GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 11:40:52 ; Search time 18939 Seconds
(without alignments)
20098.352 Million cell updates/sec

Title: US-08-731-499-9_COPY_1_10000
Perfect score: 10000
Sequence: 1 CCATCATATTTCTTATTTTT.....ATTCTTAGATACGGCAGTGG 10000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	761	7.6	841	6	CD644465	AGENCOURT
	2	752	7.5	790	8	BZ604860	WH3AE21TR
	3	742.4	7.4	744	8	BZ609651	WHADE76TF
	4	727.8	7.3	2331	3	AK030249	Mus muscu
	5	648.6	6.5	690	1	AL040111	DKFZp434C
	6	640	6.4	640	4	BM834501	K-EST0109
	7	618	6.2	618	5	BX955300	DKFZp781N
	8	592.4	5.9	595	4	BF944766	PM3-NN117
	9	581.4	5.8	583	5	BP336988	BP336988
	10	557.6	5.6	775	8	BZ603781	WH3AE51TR
	11	543.8	5.4	571	7	CR753532	DKFZp469L
	12	536	5.4	585	5	BP366185	BP366185
c	13	528.2	5.3	557	8	B16035	346L18.TP C
	14	516	5.2	598	5	BP270262	BP270262
	15	498.4	5.0	583	5	BP229047	BP229047
	16	485.8	4.9	730	8	BZ606973	WH3AE02TR
c	17	473.8	4.7	569	2	AW753609	RC1-CT026
	18	473.2	4.7	895	9	CR094054	Forward s
	19	418.8	4.2	768	2	BF237807	601842326
	20	402.8	4.0	458	5	BP201868	BP201868
	21	394	3.9	411	4	BG007233	IL5-GN023
c	22	382	3.8	479	1	AI380386	tf96a01.x
c	23	379.8	3.8	594	7	CV409050	RC1-HT026
	24	378.2	3.8	636	6	CB457791	CB457791 715904 MA
c	25	373.2	3.7	400	5	BQ367562	BQ367562 IL5-GN023
	26	369	3.7	399	2	BE819616	BE819616 RC2-BN033
	27	365.4	3.7	368	4	BM146822	BM146822 TCAAP1E74
c	28	352.4	3.5	381	2	BF083401	BF083401 RC6-CT043
	29	350.8	3.5	412	1	AA069107	AA069107 zm10b10.r
c	30	349	3.5	703	4	BI185557	BI185557 UNL-P-FN-
c	31	348.8	3.5	353	4	BG013874	BG013874 IL5-GN023
c	32	346.6	3.5	662	6	CB460210	CB460210 720064 MA
c	33	341.6	3.4	440	7	N70546	N70546 za83c12.s1
	34	316.4	3.2	674	6	BY748714	BY748714 BY748714
	35	308.8	3.1	320	8	B41669	B41669 HS-1054-A1-
c	36	307.8	3.1	667	8	AZ939266	AZ939266 2M0198I12
	37	306	3.1	307	8	AZ537566	AZ537566 AST-2P596
	38	302.4	3.0	352	4	BF963063	BF963063 PM3-NN117
c	39	301.8	3.0	305	4	BG982616	BG982616 IL5-CN006
	40	299	3.0	328	2	BE818091	BE818091 CM2-BN027
	41	282.6	2.8	576	2	BF782572	BF782572 602106595
c	42	282	2.8	327	2	BE817937	BE817937 CM2-BN027
	43	272	2.7	651	5	BX119280	BX119280 BX119280
c	44	270.2	2.7	482	8	AQ133619	AQ133619 HS_3047_A
c	45	269	2.7	511	8	AQ512276	AQ512276 HS_5077_A

ALIGNMENTS

RESULT 1
CD644465
LOCUS CD644465 841 bp mRNA linear EST 17-JUN-2003
DEFINITION AGENCOURT_14556793 NIA Human H1 Embryonic Stem Cell cDNA Library (Long) Homo sapiens cDNA clone IMAGE:30427829 5', mRNA sequence.
ACCESSION CD644465
VERSION CD644465.1 GI:31816375
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 841)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
cDNA Library Preparation: Yulan Piao and Minoru Ko
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC c lone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: NDAM515 row: o column: 06
High quality sequence stop: 667.
FEATURES Location/Qualifiers
source 1. .841
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30427829"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
/note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line . Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, Wi, cultured according to their instructions, on

MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with TRIzol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199]) Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The average insert size is about 3.6kb."

ORIGIN

Query Match 7.6%; Score 761; DB 6; Length 841;
Best Local Similarity 99.1%; Pred. No. 2.8e-110;
Matches 775; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy       9063 TACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGTAA 9122
             ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db          1 TACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCTCTAGTTTCTGTAA 60

Qy       9123 ACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAA 9182
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAA 120

Qy       9183 GCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACTTT 9242
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACTTT 180

Qy       9243 AGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCCGC 9302
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCCGC 240

Qy       9303 CACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGATAA 9362
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGATAA 300

Qy       9363 GACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACCCT 9422
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACCCT 360

Qy       9423 CGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGGGC 9482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGGGC 420

Qy       9483 ACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTTGG 9542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTTGG 480

Qy       9543 TGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCCGG 9602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCCGG 540

Qy       9603 GGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATC 9662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATC 600

Qy       9663 ACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTCCT 9722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTCCT 660

Qy       9723 GAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGCTC 9782
             |||||||||||||||||||||||| |||||||||||||||||||  ||||||||||||
Db        661 GAGCTCCAGCGAGGTCGATTCTCC-AATGTGCTGACTGTTCAGAAAGCCTATGGTGGCTC 719

Qy       9783 CGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGG 9842
             ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        720 CGGGCCACTTTACANCTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGG 779

Qy       9843 TA 9844
              |
Db        780 AA 781
```

RESULT 2
BZ604860

```
LOCUS       BZ604860                790 bp    DNA     linear   GSS 08-JUN-2003
DEFINITION  WH3AE21TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
            sapiens genomic clone MCF7_1-26E17, genomic survey sequence.
ACCESSION   BZ604860
VERSION     BZ604860.1  GI:31513368
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 790)
  AUTHORS   Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
            Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
            Gray,J.W. and Collins,C.
  TITLE     End-sequence profiling: Sequence-based analysis of aberrant genomes
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
  MEDLINE   22709111
   PUBMED   12788976
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .790
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_1-26E17"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."
ORIGIN
```

Query Match 7.5%; Score 752; DB 8; Length 790;
Best Local Similarity 99.2%; Pred. No. 7.5e-109;
Matches 788; Conservative 0; Mismatches 0; Indels 6; Gaps 3;

```
Qy       3895 TCAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCG 3954
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCG 60

Qy       3955 GCGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACC 4014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACC 120

Qy       4015 TCGCCGCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACG 4074
```

```
Db        121 TCGCCGCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACG 180

Qy       4075 GATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCG 4134

Db        181 GATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCG 240

Qy       4135 TGCTGTTCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGA 4194

Db        241 TGCTGTTCCGCCTG----TGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGA 296

Qy       4195 CAACGCTGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTAT 4254

Db        297 CAACGCTGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTAT 356

Qy       4255 CACTCTGTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAG 4314

Db        357 CACTCTGTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAG 416

Qy       4315 GGTTGAGTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCG 4374

Db        417 GGTTGAGTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCG 476

Qy       4375 TTTTGCTCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACT 4434

Db        477 TTTTGCTCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACT 536

Qy       4435 GCTTCTTAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTT 4494

Db        537 GCTTCTTAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTT 596

Qy       4495 GGTTGTCATACTGT-GTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGA 4553

Db        597 GGTTGTCATACTGTGGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGA 656

Qy       4554 CACTGCTCAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCC 4613

Db        657 CACTGCTCAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCC 716

Qy       4614 AAATGTCAATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTC 4673

Db        717 AAATGTCAATAGTTTGAGCATTGAGAGA-CCTAGCCTTCACTTAAGTTTTTCTGGCGTTC 775

Qy       4674 CTGATCTTTTTCTG 4687

Db        776 CTGATCTTTTTCTG 789
```

```
RESULT 3
BZ609651
LOCUS       BZ609651                744 bp    DNA     linear   GSS 08-JUN-2003
DEFINITION  WHADB76TF Human MCF7 breast cancer cell line library (MCF7_1) Homo
            sapiens genomic clone MCF7_1-20M7, genomic survey sequence.
ACCESSION   BZ609651
VERSION     BZ609651.1  GI:31518212
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 744)
  AUTHORS   Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
            Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
            Gray,J.W. and Collins,C.
  TITLE     End-sequence profiling: Sequence-based analysis of aberrant genomes
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
  MEDLINE   22709111
   PUBMED   12788976
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_1-20M7"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."
ORIGIN
```

```
  Query Match             7.4%;  Score 742.4;  DB 8;  Length 744;
  Best Local Similarity   99.9%;  Pred. No. 2.5e-107;
  Matches  743;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        763 AGCTTCTATGTATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATAT 822

Db          1 AGCTTCTATGTATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATAT 60

Qy        823 TCTTTTCAATGTGTAGCTCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATC 882

Db         61 TCTTTTCAATGTGTAGCTCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATC 120

Qy        883 CTATAGTTTAACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACA 942

Db        121 CTATAGTTTAACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACA 180

Qy        943 ACTTGAAGGTTGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCT 1002

Db        181 ACTTGAAGGTTGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCT 240

Qy       1003 GGTCACCAGCCCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGAC 1062

Db        241 GGTCACCAGCCCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGAC 300

Qy       1063 ATTCTTTCCTTTACCCTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAA 1122

Db        301 ATTCTTTCCTTTACCCTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAA 360

Qy       1123 GATCTAGTTTATTCAAAACTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGT 1182

Db        361 GATCTAGTTTATTCAAAACTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGT 420

Qy       1183 AAAGGAGGAGATGTTTCCTTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGA 1242

Db        421 AAAGGAGGAGATGTTTCCTTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGA 480

Qy       1243 AGTAGAGAGAGGGAGGAAAGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAG 1302

Db        481 AGTAGAGAGAGGGAGGAAAGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAG 540

Qy       1303 TAGTTTGCATGCCTTTTGGTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTT 1362

Db        541 TAGTTTGCATGCCTTTTGGTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTT 600

Qy       1363 GGGACTATTTAAAAACAATGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTT 1422

Db        601 GGGACTATTTAAAAACAATGGCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTT 660

Qy       1423 TAGAAGTTAATTCTTCTCCAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTA 1482

Db        661 TAGAAGTTAATTCTTCTCCAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTA 720

Qy       1483 AATAAAGTAATTTTCCACCTGGCC 1506
```

```
                    ||||||||||||||||||||||||
Db        721 AATAAAGTAATTTTCCACCTGGCC 744


RESULT 4
AK030249
LOCUS       AK030249               2331 bp    mRNA    linear   HTC 03-APR-2004
DEFINITION  Mus musculus adult male testis cDNA, RIKEN full-length enriched
            library, clone:4933431C08 product:weakly similar to ZINC FINGER
            PROTEIN 217 [Homo sapiens], full insert sequence.
ACCESSION   AK030249
VERSION     AK030249.1  GI:26081731
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 2331)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.jp/
            URL:http://fantom.gsc.riken.jp/.
FEATURES             Location/Qualifiers
     source          1. .2331
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:4933431C08"
                     /db_xref="taxon:10090"
                     /clone="4933431C08"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     misc_feature    228. .2104
                     /note="putative
                     weakly similar to ZINC FINGER PROTEIN 217 [Homo sapiens]
                     (SWISSPROT|O75362, evidence: FASTY, 69.5%ID, 67.8%length,
                     match=2090)"
ORIGIN

  Query Match             7.3%;  Score 727.8;  DB 3;  Length 2331;
  Best Local Similarity   70.2%;  Pred. No. 4e-105;
  Matches 1099;  Conservative    0;  Mismatches  412;  Indels   54;  Gaps    7;

Qy       8286 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACATCTCT 8345
              ||||||||||||||||||||||||||| ||||| ||||| ||||||||||||||||||||
Db        481 AGGTGAAAAACCATACAAATGTGAATTCTGTGAGTATGCCGCAGCCCAGAAGACATCTCT 540

Qy       8346 GAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAAC---CGATGTTGCTGCTGAAGT 8402
              |||||| |||||||||||||||||||| || || ||  |    |||| ||| |||||
Db        541 GAGGTACCACTTGGAGAGACATCACAAAGACAAGCAGCCGGTGGATGCTGCCGCTGAGTC 600

Qy       8403 CAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAA---CCGCTGACAGTGC 8459
              ||| |  || ||    | ||||  |  | ||||| ||| |||    | ||||||||||
Db        601 CAAAAGTGAAGGCCGGAGCCAGGAGCCGCAGGATGCGCTACTAACGGCTGCTGACAGTGC 660

Qy       8460 GCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACC 8519
              ||| |||||||||| || |||||| ||||||||||||||||||||  || || |||||
Db        661 GCAGACCAAAAATTTAAAGAGATTTCTTGATGGTGCCAAAGATGTTAAGGGAAGCCCACC 720

Qy       8520 TGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCT 8579
              ||| |||||||||||||||||||||||| || |||| |||||            ||
Db        721 TGCCAAGCAGCTTAAGGAGATGCCTTCTGTCTTCCAGAGTGTTC------------TCTC 768

Qy       8580 CTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGA 8639
                || ||     ||| |||||||||||||||||||| ||||||||||   ||||||||
Db        769 ACCAGCACACAGCAACGATACTCAGGATTTCCATAAACATGCAGCTGAT---AGTGCTGA 825

Qy       8640 TAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGT 8699
              |||| ||  || | |||| ||||| |||| |||||| ||   | |||||     |||||
Db        826 GAAAGCGAGGAAGAGCCCTGCCCCTACTTATCTGGACATGCAGAGAAAGA---AAGCAGG 882

Qy       8700 TGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGA 8759
              ||  ||||||| |  | ||   |||| |||     ||| || ||   |      | | |||
Db        883 GGAGCCTCAGGCCAGCAGCCCTGTCTGCAGACTAGAGGGGGTTGGGTCCTTAGCACGGGA 942

Qy       8760 TGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGC 8819
              | | |                          | |   | || | | |||    |  ||
Db        943 AGCTGG--------------------------CCATAGGGAGAAGATGGATCAGGATGC 975

Qy       8820 TGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGG 8879
              ||||| |||| | | |||  |||  || || || || |  ||||| ||| ||||| | ||
Db        976 TGACTACAGACATAAGCCCGGTGCTGACTGCCAGGACAGGCCTTTGAATCTATCCCTTGG 1035
```

```
Qy       8880 GGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTG 8939
Db       1036 GCCGCTCCACGCCTGTCCTGCAATCTCTTTGAGCAAGTGTCTGATCCCCAGCATTGCCTG 1095

Qy       8940 TCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGA 8999
Db       1096 CCCCTTTTGTACTTTCAAGACCTTTTATCCGGAAGTCCTTATGATGCACCAGAGACTTGA 1155

Qy       9000 GCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCG 9059
Db       1156 GCACAGGTACAACCCTGACCCGCACAAGAACGGCAGCAGCAAGTCTGTGCTGAGGAACAG 1215

Qy       9060 ACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTG 9119
Db       1216 GCGTACCGGGTGCCCTCCGGCTTTGCTGGGAAAAGATGTCCCTCCCTTGTCTGGCCTGCA 1275

Qy       9120 TAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGG 9179
Db       1276 CAAGCCCAAGGCCAAGACTGCCTTCTCACCACACTCGAAGTCCCTGCACTCAGAGAAGGC 1335

Qy       9180 GAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCAC 9239
Db       1336 TCGGCAGGGGGCCTCGGGGCCAAGCAAAGCACCCCAGACTTCAGGACCAGACAACAGCAC 1395

Qy       9240 TTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGC 9299
Db       1396 TTTAGCCCCAAGTAACCTGAAGTCACACAGGTCACAACCCAATGCTGGGGGCACCAG--- 1452

Qy       9300 CGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGA 9359
Db       1453 TGCCACCAGGCAGCAGCAGTCAGAGTTGTTTCCCAAAGGTGGCGTCCCTGCTGCTATGGA 1512

Qy       9360 TAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCAC 9419
Db       1513 TAAGGTGAAGAGACCTGAGCCAAAACTGAAGTCCCTACCAGCCTCCCCGTCTCAGTCCCC 1572

Qy       9420 CCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTG 9479
Db       1573 CCTCAGCAGTAATAATAGCAACGGTTCTGTTGAGTATCCCGTGAAGGTTGACGGCCCATG 1632

Qy       9480 GGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATT 9539
Db       1633 GGCACAGCAAGGGAGAGACTACTACTGCCATCGGAATTCTGGCAGTGCCGCAGCAGAGTA 1692

Qy       9540 TGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCC 9599
Db       1693 CAGTGAGCCACATCCCAAAAGACTCAAGTCCAGTGCGGTGTCCCTGGACACAGAGCATGC 1752

Qy       9600 CGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCAC 9659
Db       1753 AGGGACCAATGGCAGAAGGGGCTTTGAGCTCCCCAAGTACCATGTGGTCAGGAGCATCAC 1812

Qy       9660 ATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTT 9719
Db       1813 CTCCTTGCTACCACCAGAGTGTGTGCGCCCACCGCCTGTGCTGCCCCACAAAGCCCGTTT 1872

Qy       9720 CCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGG 9779
Db       1873 CCTGAGCCCTGGGGAGGTGGAGTCACCCAGTGTGTTGGCTGTGCAGAAGCCCTACAGTGC 1932

Qy       9780 CTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGA 9839
Db       1933 CTCTGGACCCCTGTATACCTGTGGACCCGTGGGACACGCAGGAGGCAGCCCAGCCCTTGA 1992

Qy       9840 AGGTA 9844
Db       1993 AGGAA 1997
```

RESULT 5
AL040111
LOCUS AL040111 690 bp mRNA linear EST 04-SEP-2003

DEFINITION DKFZp434C2113_r1 434 (synonym: htes3) Homo sapiens cDNA clone DKFZp434C2113 5', mRNA sequence.
ACCESSION AL040111
VERSION AL040111.1 GI:5409078
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 690)
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE EST (Koehrer, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
s1 sequence also available.
This clone (DKFZp434C2113) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES Location/Qualifiers
source 1. .690
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434C2113"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
ORIGIN

Query Match 6.5%; Score 648.6; DB 1; Length 690;
Best Local Similarity 98.0%; Pred. No. 1.7e-92;
Matches 676; Conservative 0; Mismatches 12; Indels 2; Gaps 2;

```
Qy       1056 TCAGGACATTCTTTCCTTTACCCTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAA 1115
Db          1 TCNGGACATTCTTTCCTTTACCCTAATCCTTGATCTTCAGTCTTATTTAGAAAAGCTTAA 60

Qy       1116 TGTTAAAGATCTAGTTTATTCAAAACTAAAGATAACAAGGAGTATGAGAATTTCTATTTC 1175
Db         61 TGTTAAAGATCTAGTTTATTCAAAACTAAAGATAACAAGGAGTATGAGAATTTCTATTTC 120

Qy       1176 GGAGTGTAAAGGAGGAGATGTTTCCTTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCT 1235
Db         121 GGAGTGTAAAGGAGGAGATGTTTCCTTGGCTTCTCTGAGCCTACAGGCCTTCCTTGCTCT 180

Qy       1236 TTAAGGAAGTAGAGAGAGGGAGGAAAGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTG 1295
Db         181 TTAAGGAAGTAGAGAGAGGGAGGAAAGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTG 240

Qy       1296 CTGGGAGTAGTTTGCATGCCTTTTGGTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTA 1355
Db         241 CTGGGAGTAGTTTGCATCCCTTTTGGTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTA 300

Qy       1356 AGTAGTTGGGACTATTTAAAAACAATGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTA 1415
Db         301 AGTAGTTGGGACTATTTAGAAACAATGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTA 360

Qy       1416 TTGGCTTTAGAAGTTAATTCTTCTCCAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGC 1475
Db         361 TTGGCTTTAGAAGTTAATTCTTCTCCAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGC 420

Qy       1476 AGAGCTAAATAAAGTAATTTTCCACCTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAA 1535
```

```
Db        421 AGAGCTAAATAAAGTAATTTTCCACCTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAA 480

Qy       1536 AATGAGAGGGCCCATTTTCTGATGAAAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGA 1595
              ||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AATCAGAGGACCCATTTTCTGATGAAAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGA 540

Qy       1596 TTTTATTTTAAATCTATACATTCACAAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTG 1655
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        541 TTTTATTTTAAATCTATACNTTCACAAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTG 600

Qy       1656 TTTGGACTAGGAAGCCCCACCCAGTGCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAG-C 1714
              |||||||||||||||||||||||||||||||||||||||||||| || |||||| || |
Db        601 TTTGGACTAGGAAGCCCCACCCAGTGCTTGTTGAAGGCAGAAAGTTGTGGAAAGCNAGCC 660

Qy       1715 TGGGATTTGAACAG-TGGATTGAGGTTTCG 1743
              ||||||||||||| |||||||||||||||
Db        661 CGGGATTTGAACAGTTGGATTGAGGTTTCG 690
```

```
RESULT 6
BM834501
LOCUS       BM834501                640 bp    mRNA    linear   EST 06-MAR-2002
DEFINITION  K-EST0109502 S11SNU1 Homo sapiens cDNA clone S11SNU1-63-D04 5',
            mRNA sequence.
ACCESSION   BM834501
VERSION     BM834501.1  GI:19190910
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 640)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 63  row: D  column: 04
            High quality sequence stop: 640.
FEATURES             Location/Qualifiers
     source          1. .640
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S11SNU1-63-D04"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Lymphoblast-like"
                     /cell_line="SNU-1"
                     /lab_host="Top10F'"
                     /clone_lib="S11SNU1"
                     /note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI;
                     Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including SfiI
                     site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized with Superscript II using SfiI
                     oligo-dT primer. After first strand synthesis, RNA was
                     degraded by NaOH treatment and cDNA was amplified by PCR
                     reaction. The PCR products were digested with SfiI and
                     cloned into DraIII- digested pME18S-FL3 vector. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
```

```
ORIGIN

  Query Match             6.4%;  Score 640;  DB 4;  Length 640;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-91;
  Matches  640;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2489 TGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGTTTGTG 2548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGTTTGTG 60

Qy       2549 GAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGGACTAT 2608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGGACTAT 120

Qy       2609 ACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTTGGAACCCAGGGCTGT 2668
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTTGGAACCCAGGGCTGT 180

Qy       2669 TAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTCCAGAAATTGCCCAAC 2728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTCCAGAAATTGCCCAAC 240

Qy       2729 GGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTT 2788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTT 300

Qy       2789 AATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGA 2848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGA 360

Qy       2849 GGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGAAAAAAA 2908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGAAAAAAA 420

Qy       2909 TGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGACCTTCAC 2968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGACCTTCAC 480

Qy       2969 ACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTGTGAACC 3028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTGTGAACC 540

Qy       3029 AGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCGAAC 3088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCGAAC 600

Qy       3089 AGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTT 3128
              |||||||||||||||||||||||||||||||||||||||
Db        601 AGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTT 640
```

```
RESULT 7
BX955300
LOCUS       BX955300                618 bp    mRNA    linear   EST 01-MAR-2004
DEFINITION  DKFZp781N0270_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
            DKFZp781N0270 5', mRNA sequence.
ACCESSION   BX955300
VERSION     BX955300.1  GI:43437635
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 618)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
```

```
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp781N0270) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .618
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp781N0270"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="781 (synonym: hlcc4)"
                     /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             6.2%;  Score 618;  DB 5;  Length 618;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-87;
  Matches  618;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2802 GGGCCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCA 2861
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCA 60

Qy       2862 ATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATC 2921
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATC 120

Qy       2922 GAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGAC 2981
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGAC 180

Qy       2982 CTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGG 3041
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGG 240

Qy       3042 GTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAG 3101
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAG 300

Qy       3102 GAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTC 3161
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTC 360

Qy       3162 GCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAAC 3221
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAAC 420

Qy       3222 ATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAAT 3281
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAAT 480

Qy       3282 GGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATC 3341
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATC 540

Qy       3342 AACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATG 3401
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATG 600

Qy       3402 GTTTGTGGCTTCCTATTT 3419
             ||||||||||||||||||
Db        601 GTTTGTGGCTTCCTATTT 618


RESULT 8
BF944766
LOCUS       BF944766                595 bp    mRNA    linear   EST 22-JAN-2001
DEFINITION  PM3-NN1174-181000-001-c08 NN1174 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF944766
VERSION     BF944766.1  GI:12362041
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 595)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM3&t2=PM3-NN1174-
            181000-001-c08&t3=2000-10-18&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 585.
FEATURES             Location/Qualifiers
     source          1. .595
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="NN1174"
                     /note="Organ: nervous_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             5.9%;  Score 592.4;  DB 4;  Length 595;
  Best Local Similarity   99.8%;  Pred. No. 1.4e-83;
  Matches  593;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3041 GGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAA 3100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAA 60

Qy       3101 GGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGT 3160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGT 120

Qy       3161 CGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAA 3220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAA 180

Qy       3221 CATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAA 3280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAA 240

Qy       3281 TGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGAT 3340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGAT 300
```

```
Qy       3341 CAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCAT 3400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCAT 360

Qy       3401 GGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACAC 3460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACAC 420

Qy       3461 CAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCC 3520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCC 480

Qy       3521 GTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAAC 3580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAAC 540

Qy       3581 GGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGTTCACCACCTTCCA 3634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        541 GGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGTTCACCACCTTCAA 594


RESULT 9
BP336988
LOCUS       BP336988                583 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP336988 Sugano cDNA library, coronary artery smooth muscle cell
            Homo sapiens cDNA clone SMR07185, mRNA sequence.
ACCESSION   BP336988
VERSION     BP336988.1  GI:52266581
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="SMR07185"
                     /tissue_type="coronary artery"
                     /cell_type="smooth muscle cell"
                     /clone_lib="Sugano cDNA library, coronary artery smooth
                     muscle cell"
ORIGIN

  Query Match             5.8%;  Score 581.4;  DB 5;  Length 583;
  Best Local Similarity   99.8%;  Pred. No. 7.7e-82;
  Matches  582;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8430 TGAAGATGCACTATTAACCGCTGACAGTGCGCAAACCAAAAATTTGAAAAGATTTTTTGA 8489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGAAGATGCACTATTAACCGCTGACAGTGCGCAAACCAAAAATTTGAAAAGATTTTTTGA 60

Qy       8490 TGGTGCCAAAGATGTTACAGGCAGTCCACCTGCAAAGCAGCTTAAGGAGATGCCTTCTGT 8549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGTGCCAAAGATGTTACAGGCAGTCCACCTGCAAAGCAGCTTAAGGAGATGCCTTCTGT 120

Qy       8550 TTTTCAGAATGTTCTGGGCAGCGCTGTCCTCTCACCAGCACACAAAGATACTCAGGATTT 8609
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTTCAGAATGTTCTGGGCAGCGCTGTCCTCTCACCAGCACACAAAGATACTCAGGATTT 180

Qy       8610 CCATAAAAATGCAGCTGATGACAGTGCTGATAAAGTGAATAAAAACCCTACCCCTGCTTA 8669
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATAAAAATGCAGCTGATGACAGTGCTGATAAAGTGAATAAAAACCCTACCCCTGCTTA 240

Qy       8670 CCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAACTCAGGCAAATAACCTCATCTGTAG 8729
             ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGACCTGTTAAAAAGGAGATCAGCAGTTGAAACTCAGGCAAATAACCTCATCTGTAG 300

Qy       8730 AACCAAGGCGGATGTTACTCCTCCTCCGGATGGCAGTACCACCCATAACCTTGAAGTTAG 8789
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACCAAGGCGGATGTTACTCCTCCTCCGGATGGCAGTACCACCCATAACCTTGAAGTTAG 360

Qy       8790 CCCCAAAGAGAAGCAAACGGAGACCGCAGCTGACTGCAGATACAGGCCAAGTGTGGATTG 8849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCCCAAAGAGAAGCAAACGGAGACCGCAGCTGACTGCAGATACAGGCCAAGTGTGGATTG 420

Qy       8850 TCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCTTCACAATTGCCCGGCAATTTCTTT 8909
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCTTCACAATTGCCCGGCAATTTCTTT 480

Qy       8910 GAGTAAAAGTTTGATTCCAAGTATCACCTGTCCATTTTGTACCTTCAAGACATTTTATCC 8969
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGTAAAAGTTTGATTCCAAGTATCACCTGTCCATTTTGTACCTTCAAGACATTTTATCC 540

Qy       8970 AGAAGTTTTAATGATGCACCAGAGACTGGAGCATAAATACAAT 9012
             |||||||||||||||||||||||||||||||||||||||||||
Db        541 AGAAGTTTTAATGATGCACCAGAGACTGGAGCATAAATACAAT 583


RESULT 10
BZ603781
LOCUS       BZ603781                775 bp    DNA     linear   GSS 08-JUN-2003
DEFINITION  WH3AE51TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
            sapiens genomic clone MCF7_1-159C23, genomic survey sequence.
ACCESSION   BZ603781
VERSION     BZ603781.1  GI:31512243
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 775)
  AUTHORS   Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
            Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
            Gray,J.W. and Collins,C.
  TITLE     End-sequence profiling: Sequence-based analysis of aberrant genomes
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
  MEDLINE   22709111
   PUBMED   12788976
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .775
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_1-159C23"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."
ORIGIN
```

```
  Query Match             5.6%;  Score 557.6;  DB 8;  Length 775;
  Best Local Similarity   99.1%;  Pred. No. 4.2e-78;
  Matches  571;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 259

Qy         61 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 319

Qy        121 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 379

Qy        181 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 439

Qy        241 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 499

Qy        301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 559

Qy        361 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 619

Qy        421 AATGTTGAGAGCTGACATAGGGACTAACTT-GGTTTTTTTTTTTTTTTTTTTTTTTCAAA 479
              ||||||||||||||||||||||||||||| | ||||||||||||||||||||||||||||
Db        620 AATGTTGAGAGCTGACATAGGGACTAACTTAGTTTTTTTTTTTTTTTTTTTTTTTTCAAA 679

Qy        480 TTCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAAAACTCCA 539
              ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db        680 TTCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAAACTCCA 739

Qy        540 CTATTGCCTATTGCCACTATTTGATTTTTTAAAAAA 575
              |||||||||||||||| ||||||||||||||||||||
Db        740 CTATTGCCTATTGCCCCTATTTGATTTTTTAAAAAA 775


RESULT 11
CR753532
LOCUS       CR753532                571 bp    mRNA    linear   EST 01-SEP-2004
DEFINITION  DKFZp469L046_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
            DKFZp469L046 5', mRNA sequence.
ACCESSION   CR753532
VERSION     CR753532.1  GI:51845942
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE   1  (bases 1 to 571)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Wambutt,R., Heubner,D., Mewes,H.W., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert. Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; sequenced by Agowa
            (Berlin/Germany) within the cDNA sequencing consortium of the
            German Genome Project. This clone (DKFZp469L046) is available at
            the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
            Berlin, Germany. Please contact RZPD for ordering:
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469L046
            Further information about the clone and the sequencing project is
            available at http://mips.gsf.de/projects/cdna/.
FEATURES              Location/Qualifiers
     source           1. .571
                      /organism="Pongo pygmaeus"
                      /mol_type="mRNA"
                      /db_xref="taxon:9600"
                      /clone="DKFZp469L046"
                      /tissue_type="kidney"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="469 (synonym: pkid1)"
                      /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

  Query Match             5.4%;  Score 543.8;  DB 7;  Length 571;
  Best Local Similarity   97.0%;  Pred. No. 6.8e-76;
  Matches  554;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       2478 AAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATG 2537
              |||||||||||||||||||||||||||||||||||||| ||||||| || | ||||||||
Db          1 AAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCAAGGCATTTGGGAGCCTGTATG 60

Qy       2538 CCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGT 2597
              || |||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db         61 CCGTGTTTGTGGAAAGAACCAGTGACACCACCACTGAGCTTCCTAAAAGTTCGAAGAAGT 120

Qy       2598 TAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTTGGAA 2657
              |||||||||||||| |||||||||||||||||||||||||||||||||||||| | | |
Db        121 TAGAGGACTATACGCTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTCGAAG 180

Qy       2658 CCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTCCAGA 2717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTCCAGA 240

Qy       2718 AATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACT 2777
              |||| |||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        241 AATTCCCCAACGGAACTTTGAGATTATATGCAATCAAAAGTGACAGGAAACATGCCAACT 300

Qy       2778 CAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCG 2837
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        301 CAGTCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAGTCAG 360

Qy       2838 ATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACA 2897
              |||||||||||||| ||| ||||||||||||||||||||||||||||||||||||||||
Db        361 ATGGAGATGGAGGACGCCATGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACG 420

Qy       2898 CAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGC 2957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGC 480

Qy       2958 CAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACC 3017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACC 540

Qy       3018 CTCTGTGAACCAGCAGTTCTTCGGGTTGAAG 3048
              |||||||||||||||||||||||||||||||
Db        541 CTCTGTGAACCAGCAGTTCTTCGGGTTGAAG 571


RESULT 12
BP366185
LOCUS       BP366185                585 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo
            sapiens cDNA clone TIR09337, mRNA sequence.
ACCESSION   BP366185
VERSION     BP366185.1  GI:52296435
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 585)
```

```
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TIR09337"
                     /tissue_type="lung"
                     /cell_type="fibroblast"
                     /cell_line="TIG"
                     /dev_stage="fetal"
                     /clone_lib="Sugano cDNA library, fetal lung fibroblast
                     TIG"
ORIGIN

  Query Match             5.4%;  Score 536;  DB 5;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-74;
  Matches  536;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         50 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 109

Qy       2461 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 169

Qy       2521 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 229

Qy       2581 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 289

Qy       2641 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 349

Qy       2701 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 409

Qy       2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        410 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 469

Qy       2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        470 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 529

Qy       2881 TTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCC 2936
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        530 TTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCC 585


RESULT 13
B16035/c
LOCUS       B16035                   557 bp    DNA     linear   GSS 04-JUN-1998
DEFINITION  346L18.TP CIT978SKA1 Homo sapiens genomic clone A-346L18, genomic
            survey sequence.
ACCESSION   B16035
VERSION     B16035.1  GI:2123784
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 557)
  AUTHORS   Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.
  TITLE     Use of a BAC End Sequence Database for Sequence-Ready Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .557
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="A-346L18"
                     /sex="Female"
                     /cell_type="Fibroblast"
                     /clone_lib="CIT978SKA1"
                     /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
                     CalTech Human BAC Library A1"
ORIGIN

  Query Match             5.3%;  Score 528.2;  DB 8;  Length 557;
  Best Local Similarity   96.8%;  Pred. No. 2e-73;
  Matches  539;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy        520 AAAAAAACCAAAAAACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAG 579
              |||||| |||||||||||||||||||| |||||||||||||||| ||||||||||||||
Db        557 AAAAAACCCAAAAAACTCCACTATTGCATATTGCCACTATTTGATCTTTTAAAAAATAAG 498

Qy        580 CGTATTTTAGCATCTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCA 639
              |||||||||||||||  |||||||||| |||||||||||||||||| ||||||||||||
Db        497 CGTATTTTAGCATCTGCAAGTAGGAAGGTCCTCAAATAAATGAGTCTGTGTTCTTGGCCA 438

Qy        640 GGGAAAACAGCGTTGTCAGAAATTTGATAACTGTTTTTTCTAGGGTATGTGCTGTTATTCAG 699
              |||||||||||||||||| || || |||||||| ||||||||||||||||||||||||||
Db        437 GGGAAAACAGCGTTGTCAGCATGTGTTAACTGTTGTTCTAGGGTATGTGCTGTTATTCAG 378

Qy        700 TTAAAACCTTGCCTGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAAC 759
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 TTAAAACCTTGCCTGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAAC 318

Qy        760 TTAAGCTTCTATGTATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAA 819
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        317 TTAAGCTTCTATGTATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAACTGAA 258

Qy        820 TATTCTTTTCAATGTGTAGCTCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCC 879
              || |||||||| ||||||| ||||||||||||||||||||||||||||||||||||||||
Db        257 TAGTCTTTTCAGTGTGTAGATCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCC 198

Qy        880 ATCCTATAGTTTAACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTG 939
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 ATCCTAGAGTTTAACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTG 138

Qy        940 ACAACTTGAAGGTTGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAA 999
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 ACAACTTGAAGGTTGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAA 78

Qy       1000 GCTGGTCACCAGCCCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAG 1059
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db         77 GCTGGTCACCAGCCCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAGAGACTTCAG 18
```

```
Qy       1060 GACATTCTTTCCTTTAC 1076
              |||| ||||||||||||
Db         17 GACACTCTTTCCTTTAC 1

RESULT 14
BP270262
LOCUS       BP270262                598 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP270262 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone KAR01734, mRNA sequence.
ACCESSION   BP270262
VERSION     BP270262.1  GI:52219609
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 598)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .598
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KAR01734"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

  Query Match             5.2%;  Score 516;  DB 5;  Length 598;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-71;
  Matches  516;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 142

Qy       2461 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 202

Qy       2521 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 262

Qy       2581 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 322

Qy       2641 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 382

Qy       2701 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 442

Qy       2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 502

Qy       2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 562

Qy       2881 TTCCATTCCGAGCTACACAAGAAAAAAATGTCATCC 2916
              |||||||||||||||||||||||||||||||||||
Db        563 TTCCATTCCGAGCTACACAAGAAAAAAATGTCATCC 598

RESULT 15
BP229047
LOCUS       BP229047                583 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP229047 Sugano cDNA library, fetal brain Homo sapiens cDNA clone
            FBR01825, mRNA sequence.
ACCESSION   BP229047
VERSION     BP229047.1  GI:52101957
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="FBR01825"
                     /tissue_type="brain"
                     /dev_stage="fetal"
                     /clone_lib="Sugano cDNA library, fetal brain"
ORIGIN

  Query Match             5.0%;  Score 498.4;  DB 5;  Length 583;
  Best Local Similarity   99.8%;  Pred. No. 1e-68;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 142

Qy       2461 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        143 CTCATTCAATGGAGACAAAGAGATCTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 202

Qy       2521 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 262

Qy       2581 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 322

Qy       2641 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 382

Qy       2701 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 442

Qy       2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 502

Qy       2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2880
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 562

Qy       2881 TTCCATTCCGAGCTACACAA 2900
                ||||||||||||||||||||
Db        563 TTCCATTCCGAGCTACACAA 582


Search completed: August  1, 2005, 01:38:23
Job time : 18949 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:54:32 ; Search time 3682 Seconds
(without alignments)
17578.372 Million cell updates/sec

Title: US-08-731-499-9_COPY_1_10000
Perfect score: 10000
Sequence: 1 CCATCATATTTCTTATTTTT.....ATTCTTAGATACGGCAGTGG 10000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	10000	100.0	10365	8	US-08-731-499-9	Sequence 9, Appli
	2	9980.6	99.8	36022	13	US-10-087-192-1708	Sequence 1708, Ap
	3	1639.4	16.4	5632	15	US-10-177-293-505	Sequence 505, App
	4	1559.6	15.6	3186	8	US-08-731-499-10	Sequence 10, Appl
	5	1559.6	15.6	5252	13	US-10-087-192-1709	Sequence 1709, Ap
	6	1387.2	13.9	1507	8	US-08-731-499-3	Sequence 3, Appli
c	7	1006.8	10.1	124990	20	US-10-684-422-156	Sequence 156, App
	8	820	8.2	26345	13	US-10-087-192-1705	Sequence 1705, Ap
	9	726.4	7.3	3016	13	US-10-087-192-1706	Sequence 1706, Ap
	10	617.2	6.2	707	13	US-10-027-632-104193	Sequence 104193,
	11	617.2	6.2	707	13	US-10-027-632-325113	Sequence 325113,
	12	617.2	6.2	707	17	US-10-027-632-104193	Sequence 104193,
	13	617.2	6.2	707	17	US-10-027-632-325113	Sequence 325113,
c	14	530	5.3	530	16	US-10-029-386-9711	Sequence 9711, Ap
	15	335.8	3.4	469	13	US-10-040-739-520	Sequence 520, App
c	16	256.8	2.6	501	16	US-10-029-386-292	Sequence 292, App
c	17	251	2.5	251	16	US-10-029-386-23411	Sequence 23411, A
c	18	233.8	2.3	172570	19	US-10-450-826-114	Sequence 114, App
	19	227.2	2.3	104245	17	US-10-160-807-4	Sequence 4, Appli
	20	227.2	2.3	104245	18	US-10-655-847-4	Sequence 4, Appli
c	21	227.2	2.3	170245	19	US-10-717-597-322	Sequence 322, App
	22	226.2	2.3	548	13	US-10-027-632-263139	Sequence 263139,
	23	226.2	2.3	548	17	US-10-027-632-263139	Sequence 263139,
	24	226.2	2.3	109977	20	US-10-719-993-6870	Sequence 6870, Ap
	25	226	2.3	261638	20	US-10-719-993-6882	Sequence 6882, Ap
	26	225.8	2.3	87467	19	US-10-741-601-5634	Sequence 5634, Ap
	27	225.8	2.3	87467	21	US-10-741-600-17624	Sequence 17624, A
	28	225	2.2	32248	9	US-09-764-860-802	Sequence 802, App
	29	225	2.2	32248	14	US-10-074-095-802	Sequence 802, App
	30	225	2.2	32248	17	US-10-212-872-802	Sequence 802, App
	31	224.2	2.2	818	13	US-10-027-632-163893	Sequence 163893,
	32	224.2	2.2	818	17	US-10-027-632-163893	Sequence 163893,
	33	224	2.2	66804	9	US-09-740-041-3	Sequence 3, Appli
	34	224	2.2	66804	16	US-10-389-967-3	Sequence 3, Appli
	35	223.6	2.2	3116	10	US-09-764-872-585	Sequence 585, App
	36	223.6	2.2	32221	9	US-09-764-847-1406	Sequence 1406, Ap
	37	223.6	2.2	32221	14	US-10-092-154-1406	Sequence 1406, Ap
	38	223.6	2.2	96595	11	US-09-997-722-262	Sequence 262, App
	39	223.4	2.2	43183	17	US-10-085-117-28	Sequence 28, Appl
	40	223.4	2.2	347001	19	US-10-319-908-16	Sequence 16, Appl
	41	223.2	2.2	370469	13	US-10-087-192-250	Sequence 250, App
	42	223	2.2	58837	9	US-09-982-091A-5	Sequence 5, Appli
c	43	222.8	2.2	590	13	US-10-027-632-93928	Sequence 93928, A
c	44	222.8	2.2	590	13	US-10-027-632-93929	Sequence 93929, A
c	45	222.8	2.2	590	17	US-10-027-632-93928	Sequence 93928, A

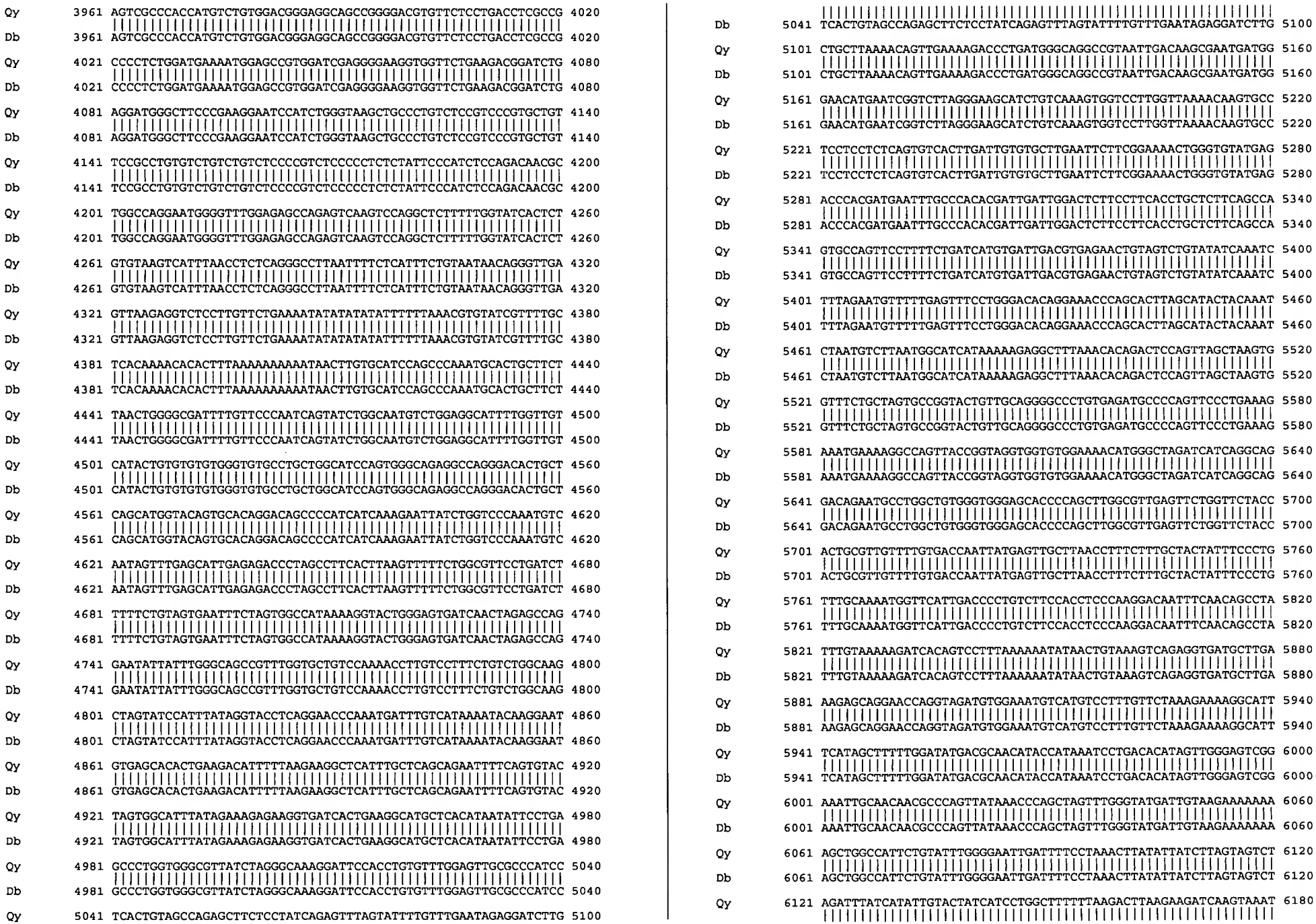
ALIGNMENTS

RESULT 1
US-08-731-499-9
; Sequence 9, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; TITLE OF INVENTION: USES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/680,395
;      FILING DATE:  15-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hunter, Tom
;      REGISTRATION NUMBER:  38,498
;      REFERENCE/DOCKET NUMBER:  2307O-068910
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10365 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1..10365
;      OTHER INFORMATION:  /note= "Genomic Sequence Encoding
;      OTHER INFORMATION:  ZABC1"
US-08-731-499-9

  Query Match             100.0%;  Score 10000;  DB 8;  Length 10365;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 10000;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 60

Qy         61 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 120

Qy        121 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 180

Qy        181 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 240

Qy        241 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300

Qy        301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 360

Qy        361 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 420

Qy        421 AATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTTCAAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTTCAAAT 480

Qy        481 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAAAACTCCAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAACCAAAAAACTCCAC 540

Qy        541 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 600

Qy        601 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 660

Qy        661 TTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 720

Qy        721 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 780

Qy        781 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 840

Qy        841 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 900

Qy        901 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 960

Qy        961 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 1020

Qy       1021 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 1080

Qy       1081 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 1140

Qy       1141 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 1200

Qy       1201 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 1260

Qy       1261 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 1320

Qy       1321 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 1380

Qy       1381 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 1440

Qy       1441 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 1500

Qy       1501 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATGA 1560

Qy       1561 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 1620

Qy       1621 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 1680

Qy       1681 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 1740

Qy       1741 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1741 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 1800
Qy       1801 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
Qy       1861 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 1920
Qy       1921 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 1980
Qy       1981 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 2040
Qy       2041 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 2100
Qy       2101 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 2160
Qy       2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 2220
Qy       2221 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 2280
Qy       2281 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 2340
Qy       2341 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTTCCTCT 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTTCCTCT 2400
Qy       2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 2460
Qy       2461 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520
Qy       2521 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2580
Qy       2581 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 2640
Qy       2641 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2700
Qy       2701 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2760
Qy       2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2820
Qy       2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2880
```

```
Qy       2881 TTCCATTCCGAGCTACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 2940
Qy       2941 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 3000
Qy       3001 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 3060
Qy       3061 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 3120
Qy       3121 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 3180
Qy       3181 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 3240
Qy       3241 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 3300
Qy       3301 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 3360
Qy       3361 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 3420
Qy       3421 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 3480
Qy       3481 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 3540
Qy       3541 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 3600
Qy       3601 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 3660
Qy       3661 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 3720
Qy       3721 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 3780
Qy       3781 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 3840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 3840
Qy       3841 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 3900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 3900
Qy       3901 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 3960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 3960
```


Qy 3961 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 4020
Db 3961 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 4020
Qy 4021 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 4080
Db 4021 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 4080
Qy 4081 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 4140
Db 4081 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 4140
Qy 4141 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 4200
Db 4141 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 4200
Qy 4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 4260
Db 4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 4260
Qy 4261 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 4320
Db 4261 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 4320
Qy 4321 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 4380
Db 4321 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 4380
Qy 4381 TCACAAAACACACTTTAAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 4440
Db 4381 TCACAAAACACACTTTAAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 4440
Qy 4441 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 4500
Db 4441 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 4500
Qy 4501 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 4560
Db 4501 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 4560
Qy 4561 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 4620
Db 4561 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 4620
Qy 4621 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 4680
Db 4621 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 4680
Qy 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATCAACTAGAGCCAG 4740
Db 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATCAACTAGAGCCAG 4740
Qy 4741 GAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGCAAG 4800
Db 4741 GAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGCAAG 4800
Qy 4801 CTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGGAAT 4860
Db 4801 CTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGGAAT 4860
Qy 4861 GTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTGTAC 4920
Db 4861 GTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTGTAC 4920
Qy 4921 TAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCCTGA 4980
Db 4921 TAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCCTGA 4980
Qy 4981 GCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCATCC 5040
Db 4981 GCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCATCC 5040
Qy 5041 TCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATCTTG 5100
Db 5041 TCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATCTTG 5100
Qy 5101 CTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGATGG 5160
Db 5101 CTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGATGG 5160
Qy 5161 GAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGTGCC 5220
Db 5161 GAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGTGCC 5220
Qy 5221 TCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTATGAG 5280
Db 5221 TCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTATGAG 5280
Qy 5281 ACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAGCCA 5340
Db 5281 ACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAGCCA 5340
Qy 5341 GTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAAATC 5400
Db 5341 GTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAAATC 5400
Qy 5401 TTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAAT 5460
Db 5401 TTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAAT 5460
Qy 5461 CTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAAGTG 5520
Db 5461 CTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAAGTG 5520
Qy 5521 GTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGAAAG 5580
Db 5521 GTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGAAAG 5580
Qy 5581 AAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGGCAG 5640
Db 5581 AAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGGCAG 5640
Qy 5641 GACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCTACC 5700
Db 5641 GACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCTACC 5700
Qy 5701 ACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCCCTG 5760
Db 5701 ACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCCCTG 5760
Qy 5761 TTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGCCTA 5820
Db 5761 TTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGCCTA 5820
Qy 5821 TTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCTTGA 5880
Db 5821 TTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCTTGA 5880
Qy 5881 AAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGCATT 5940
Db 5881 AAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGCATT 5940
Qy 5941 TCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGTCGG 6000
Db 5941 TCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGTCGG 6000
Qy 6001 AAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAAAAA 6060
Db 6001 AAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAAAAA 6060
Qy 6061 AGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAGTCT 6120
Db 6061 AGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAGTCT 6120
Qy 6121 AGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTAAAT 6180

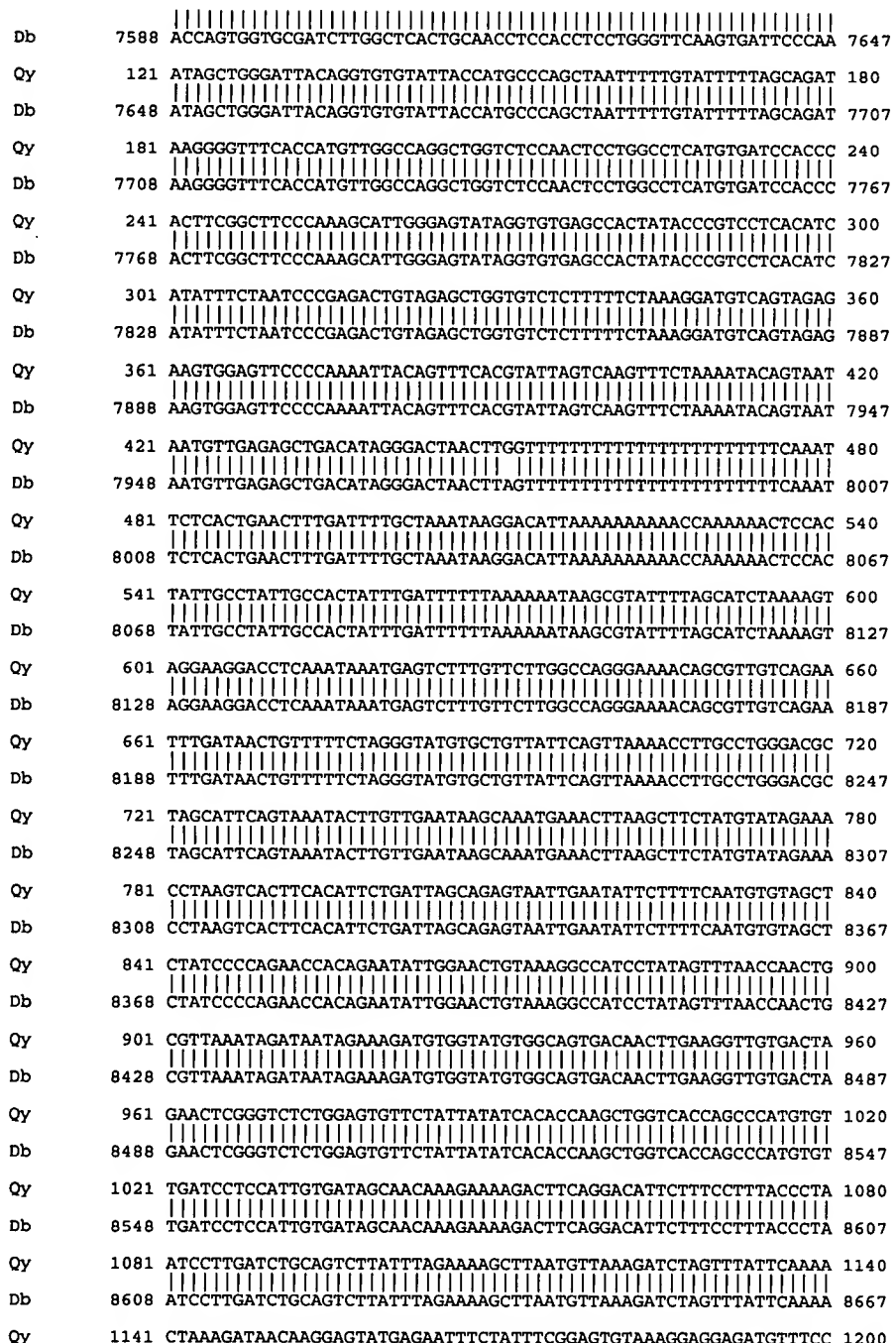
```
Db       6121 AGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTAAAT 6180
Qy       6181 TTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTGGAT 6240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6181 TTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTGGAT 6240

Qy       6241 AGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTTTGG 6300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6241 AGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTTTGG 6300

Qy       6301 GGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTTGAA 6360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6301 GGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTTGAA 6360

Qy       6361 GTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGTCCT 6420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6361 GTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGTCCT 6420

Qy       6421 ATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTGTCAG 6480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6421 ATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTGTCAG 6480

Qy       6481 ATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTGCTA 6540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6481 ATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTGCTA 6540

Qy       6541 AATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAAGAT 6600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6541 AATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAAGAT 6600

Qy       6601 TAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCTTAT 6660
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6601 TAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCTTAT 6660

Qy       6661 TTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTATTC 6720
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6661 TTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTATTC 6720

Qy       6721 TGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCTGCC 6780
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6721 TGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCTGCC 6780

Qy       6781 TGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATCTTT 6840
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6781 TGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATCTTT 6840

Qy       6841 CTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCAGAA 6900
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6841 CTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCAGAA 6900

Qy       6901 AAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGCATT 6960
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6901 AAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGCATT 6960

Qy       6961 TGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCATGCT 7020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6961 TGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCATGCT 7020

Qy       7021 TTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTCAGC 7080
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7021 TTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTCAGC 7080

Qy       7081 TACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGAAAA 7140
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7081 TACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGAAAA 7140

Qy       7141 ATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCAAAT 7200
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7141 ATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCAAAT 7200

Qy       7201 TATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAACCA 7260
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7201 TATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAACCA 7260

Qy       7261 TGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTCTGC 7320
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7261 TGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTCTGC 7320

Qy       7321 CAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATTTTT 7380
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7321 CAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATTTTT 7380

Qy       7381 ATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTTTCT 7440
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7381 ATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTTTCT 7440

Qy       7441 TTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTGGTG 7500
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7441 TTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTGGTG 7500

Qy       7501 CAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAGCCT 7560
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7501 CAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAGCCT 7560

Qy       7561 CCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTTTAG 7620
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7561 CCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTTTAG 7620

Qy       7621 TAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATC 7680
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7621 TAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATC 7680

Qy       7681 CGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTGAGC 7740
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7681 CGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTGAGC 7740

Qy       7741 CACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTTATC 7800
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7741 CACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTTATC 7800

Qy       7801 TTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAATGAA 7860
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7801 TTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAATGAA 7860

Qy       7861 ATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTAGGA 7920
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7861 ATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTAGGA 7920

Qy       7921 GGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCAAGT 7980
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7921 GGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCAAGT 7980

Qy       7981 TATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATATTTC 8040
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7981 TATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATATTTC 8040

Qy       8041 CCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATTGAA 8100
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8041 CCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATTGAA 8100

Qy       8101 GATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGAGAA 8160
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8101 GATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGAGAA 8160

Qy       8161 ATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAATAA 8220
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8161 ATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAATAA 8220

Qy       8221 ATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTTCTT 8280
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8221 ATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTTCTT 8280

Qy       8281 GCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACA 8340
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8281 GCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACA 8340
```

```
Qy       8341 TCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAA 8400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8341 TCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAA 8400

Qy       8401 GTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCG 8460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8401 GTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCG 8460

Qy       8461 CAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCT 8520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8461 CAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCT 8520

Qy       8521 GCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTC 8580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8521 GCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTC 8580

Qy       8581 TCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGAT 8640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8581 TCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGAT 8640

Qy       8641 AAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTT 8700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8641 AAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTT 8700

Qy       8701 GAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGAT 8760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8701 GAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGAT 8760

Qy       8761 GGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCT 8820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8761 GGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCT 8820

Qy       8821 GACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGG 8880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8821 GACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGG 8880

Qy       8881 GCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGT 8940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8881 GCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGT 8940

Qy       8941 CCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAG 9000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8941 CCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAG 9000

Qy       9001 CATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGA 9060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9001 CATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGA 9060

Qy       9061 CGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGT 9120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9061 CGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGT 9120

Qy       9121 AAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGG 9180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9121 AAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGG 9180

Qy       9181 AAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACT 9240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9181 AAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACT 9240

Qy       9241 TTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCC 9300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9241 TTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCC 9300

Qy       9301 GCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGAT 9360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9301 GCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGAT 9360

Qy       9361 AAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACC 9420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9361 AAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACC 9420

Qy       9421 CTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGG 9480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9421 CTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGG 9480

Qy       9481 GCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTT 9540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9481 GCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTT 9540

Qy       9541 GGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCC 9600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9541 GGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCC 9600

Qy       9601 GGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACA 9660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9601 GGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACA 9660

Qy       9661 TCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTC 9720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9661 TCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTC 9720

Qy       9721 CTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGC 9780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9721 CTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGC 9780

Qy       9781 TCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAA 9840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9781 TCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAA 9840

Qy       9841 GGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAATCCA 9900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9841 GGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAATCCA 9900

Qy       9901 GGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGTTCT 9960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9901 GGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGTTCT 9960

Qy       9961 TGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000
             ||||||||||||||||||||||||||||||||||||||||
Db       9961 TGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000
```

RESULT 2
US-10-087-192-1708
; Sequence 1708, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1708
; LENGTH: 36022
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1708

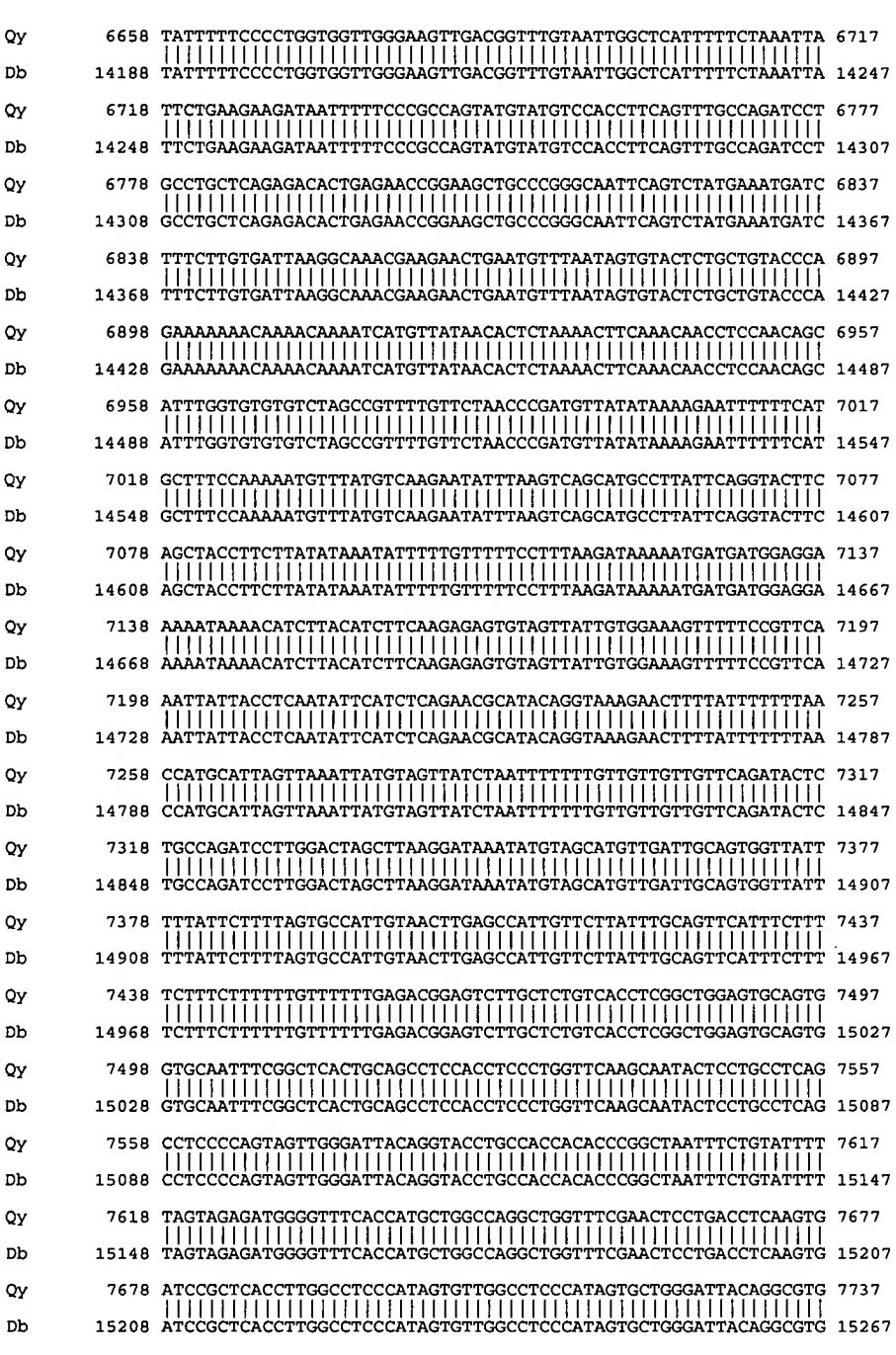
Query Match 99.8%; Score 9980.6; DB 13; Length 36022;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 9996; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

```
Qy          1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7528 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 7587

Qy         61 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 120
```


Db 7588 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 7647
Qy 121 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 180
Db 7648 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 7707
Qy 181 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 240
Db 7708 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 7767
Qy 241 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300
Db 7768 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 7827
Qy 301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 360
Db 7828 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 7887
Qy 361 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 420
Db 7888 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 7947
Qy 421 AATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTTCAAAT 480
Db 7948 AATGTTGAGAGCTGACATAGGGACTAACTTAGTTTTTTTTTTTTTTTTTTTTTTTCAAAT 8007
Qy 481 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAAAACTCCAC 540
Db 8008 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAACCAAAAAACTCCAC 8067
Qy 541 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 600
Db 8068 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 8127
Qy 601 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 660
Db 8128 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 8187
Qy 661 TTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 720
Db 8188 TTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 8247
Qy 721 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 780
Db 8248 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 8307
Qy 781 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 840
Db 8308 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 8367
Qy 841 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 900
Db 8368 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 8427
Qy 901 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 960
Db 8428 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 8487
Qy 961 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 1020
Db 8488 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 8547
Qy 1021 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 1080
Db 8548 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 8607
Qy 1081 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 1140
Db 8608 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 8667
Qy 1141 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 1200
Db 8668 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 8727
Qy 1201 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 1260
Db 8728 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 8787
Qy 1261 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 1320
Db 8788 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 8847
Qy 1321 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 1380
Db 8848 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 8907
Qy 1381 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 1440
Db 8908 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 8967
Qy 1441 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 1500
Db 8968 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 9027
Qy 1501 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAAATGAGAGGGCCCATTTTCTGATGA 1560
Db 9028 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAAATGAGAGGGCCCATTTTCTGATGA 9087
Qy 1561 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 1620
Db 9088 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 9147
Qy 1621 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 1680
Db 9148 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 9207
Qy 1681 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 1740
Db 9208 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 9267
Qy 1741 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 1800
Db 9268 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 9327
Qy 1801 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
Db 9328 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 9387
Qy 1861 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 1920
Db 9388 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 9447
Qy 1921 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 1980
Db 9448 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 9507
Qy 1981 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 2040
Db 9508 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 9567
Qy 2041 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 2100
Db 9568 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 9627
Qy 2101 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 2160
Db 9628 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 9687
Qy 2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 2220
Db 9688 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 9747
Qy 2221 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 2280
Db 9748 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 9807

```
Qy      2281 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9808 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 9867

Qy      2341 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTCT 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9868 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTCT 9927

Qy      2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9928 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 9987

Qy      2461 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9988 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 10047

Qy      2521 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10048 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 10107

Qy      2581 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10108 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 10167

Qy      2641 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10168 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 10227

Qy      2701 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10228 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 10287

Qy      2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10288 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 10347

Qy      2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10348 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 10407

Qy      2881 TTCCATTCCGAGCTACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10408 TTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 10467

Qy      2941 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10468 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 10527

Qy      3001 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10528 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 10587

Qy      3061 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10588 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 10647

Qy      3121 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10648 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 10707

Qy      3181 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10708 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 10767

Qy      3241 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10768 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 10827

Qy      3301 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10828 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 10887

Qy      3361 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10888 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 10947

Qy      3421 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10948 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 11007

Qy      3481 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11008 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 11067

Qy      3541 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11068 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 11127

Qy      3601 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11128 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 11187

Qy      3661 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11188 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 11247

Qy      3721 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11248 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 11307

Qy      3781 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 3840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11308 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 11367

Qy      3841 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 3900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11368 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 11427

Qy      3901 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 3960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11428 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 11487

Qy      3961 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 4020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11488 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 11547

Qy      4021 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 4080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11548 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 11607

Qy      4081 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 4140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11608 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 11667

Qy      4141 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 4200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11668 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 11727

Qy      4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 4260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11728 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 11787

Qy      4261 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 4320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11788 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 11847

Qy      4321 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 4380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11848 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 11907

Qy      4381 TCACAAAACACACTTTAAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 4440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11908 TCACAAAACACACTTTAAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 11967

Qy      4441 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 4500
```

```
Db      11968 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 12027
Qy       4501 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 4560
Db      12028 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 12087
Qy       4561 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 4620
Db      12088 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 12147
Qy       4621 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 4680
Db      12148 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 12207
Qy       4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAG---TGATCAACTAGAGC 4737
Db      12208 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATGATCAACTAGAGC 12267
Qy       4738 CAGGAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGC 4797
Db      12268 CAGGAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGC 12327
Qy       4798 AAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGG 4857
Db      12328 AAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGG 12387
Qy       4858 AATGTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTG 4917
Db      12388 AATGTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTG 12447
Qy       4918 TACTAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCC 4977
Db      12448 TACTAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCC 12507
Qy       4978 TGAGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCA 5037
Db      12508 TGAGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCA 12567
Qy       5038 TCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATC 5097
Db      12568 TCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATC 12627
Qy       5098 TTGCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGA 5157
Db      12628 TTGCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGA 12687
Qy       5158 TGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGT 5217
Db      12688 TGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTCGGTTAAAACAAGT 12747
Qy       5218 GCCTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTAT 5277
Db      12748 GCCTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTAT 12807
Qy       5278 GAGACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAG 5337
Db      12808 GAGACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAG 12867
Qy       5338 CCAGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAA 5397
Db      12868 CCAGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAA 12927
Qy       5398 ATCTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACA 5457
Db      12928 ATCTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACA 12987
Qy       5458 AATCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAA 5517
Db      12988 AATCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAA 13047
Qy       5518 GTGGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGA 5577
Db      13048 GTGGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGA 13107
Qy       5578 AAGAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGG 5637
Db      13108 AAGAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGG 13167
Qy       5638 CAGGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCT 5697
Db      13168 CAGGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCT 13227
Qy       5698 ACCACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCC 5757
Db      13228 ACCACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCC 13287
Qy       5758 CTGTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGC 5817
Db      13288 CTGTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGC 13347
Qy       5818 CTATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCT 5877
Db      13348 CTATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCT 13407
Qy       5878 TGAAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGC 5937
Db      13408 TGAAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGC 13467
Qy       5938 ATTTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGT 5997
Db      13468 ATTTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGT 13527
Qy       5998 CGGAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAA 6057
Db      13528 CGGAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAA 13587
Qy       6058 AAAAGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAG 6117
Db      13588 AAAAGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAG 13647
Qy       6118 TCTAGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTA 6177
Db      13648 TCTAGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTA 13707
Qy       6178 AATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTG 6237
Db      13708 AATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTG 13767
Qy       6238 GATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTT 6297
Db      13768 GATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTT 13827
Qy       6298 TGGGGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTT 6357
Db      13828 TGGGGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTT 13887
Qy       6358 GAAGTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGT 6417
Db      13888 GAAGTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGT 13947
Qy       6418 CCTATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTGT 6477
Db      13948 CCTATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATTTGGGGACAGGTTTTCATTTGT 14007
Qy       6478 CAGATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTG 6537
Db      14008 CAGATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTG 14067
Qy       6538 CTAAATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAA 6597
Db      14068 CTAAATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAA 14127
Qy       6598 GATTAAGTTGTGCTTGCTGCCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCT 6657
Db      14128 GATTAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCT 14187
```


Qy 6658 TATTTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTA 6717
Db 14188 TATTTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTA 14247
Qy 6718 TTCTGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCT 6777
Db 14248 TTCTGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCT 14307
Qy 6778 GCCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATC 6837
Db 14308 GCCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATC 14367
Qy 6838 TTTCTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCA 6897
Db 14368 TTTCTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCA 14427
Qy 6898 GAAAAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGC 6957
Db 14428 GAAAAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGC 14487
Qy 6958 ATTTGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCAT 7017
Db 14488 ATTTGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCAT 14547
Qy 7018 GCTTTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTC 7077
Db 14548 GCTTTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTC 14607
Qy 7078 AGCTACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGA 7137
Db 14608 AGCTACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGA 14667
Qy 7138 AAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCA 7197
Db 14668 AAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCA 14727
Qy 7198 AATTATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAA 7257
Db 14728 AATTATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAA 14787
Qy 7258 CCATGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTC 7317
Db 14788 CCATGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTC 14847
Qy 7318 TGCCAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATT 7377
Db 14848 TGCCAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATT 14907
Qy 7378 TTTATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTT 7437
Db 14908 TTTATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTT 14967
Qy 7438 TCTTTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTG 7497
Db 14968 TCTTTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTG 15027
Qy 7498 GTGCAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAG 7557
Db 15028 GTGCAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAG 15087
Qy 7558 CCTCCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTT 7617
Db 15088 CCTCCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTT 15147
Qy 7618 TAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTG 7677
Db 15148 TAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTG 15207
Qy 7678 ATCCGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTG 7737
Db 15208 ATCCGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTG 15267
Qy 7738 AGCCACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTT 7797
Db 15268 AGCCACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTT 15327
Qy 7798 ATCTTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAAT 7857
Db 15328 ATCTTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAAT 15387
Qy 7858 GAAATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTA 7917
Db 15388 GAAATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTA 15447
Qy 7918 GGAGGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCA 7977
Db 15448 GGAGGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCA 15507
Qy 7978 AGTTATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATAT 8037
Db 15508 AGTTATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATAT 15567
Qy 8038 TTCCCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATT 8097
Db 15568 TTCCCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATT 15627
Qy 8098 GAAGATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGA 8157
Db 15628 GAAGATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGA 15687
Qy 8158 GAAATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAA 8217
Db 15688 GAAATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAA 15747
Qy 8218 TAAATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTT 8277
Db 15748 TAAATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTT 15807
Qy 8278 CTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAG 8337
Db 15808 CTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAG 15867
Qy 8338 ACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCT 8397
Db 15868 ACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCT 15927
Qy 8398 GAAGTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGT 8457
Db 15928 GAAGTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGT 15987
Qy 8458 GCGCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCA 8517
Db 15988 GCGCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCA 16047
Qy 8518 CCTGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTC 8577
Db 16048 CCTGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTC 16107
Qy 8578 CTCTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCT 8637
Db 16108 CTCTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCT 16167
Qy 8638 GATAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCA 8697
Db 16168 GATAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCA 16227
Qy 8698 GTTGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCG 8757
Db 16228 GTTGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCG 16287
Qy 8758 GATGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCA 8817
Db 16288 GATGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCA 16347
Qy 8818 GCTGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTG 8877

```
Db      16348 GCTGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTG 16407

Qy       8878 GGGGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACC 8937

Db      16408 GGGGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACC 16467

Qy       8938 TGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTG 8997

Db      16468 TGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTG 16527

Qy       8998 GAGCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGT 9057

Db      16528 GAGCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGT 16587

Qy       9058 CGACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTC 9117

Db      16588 CGACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCTCTAGTTTC 16647

Qy       9118 TGTAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAG 9177

Db      16648 TGTAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAG 16707

Qy       9178 GGGAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGC 9237

Db      16708 GGGAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGC 16767

Qy       9238 ACTTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGG 9297

Db      16768 ACTTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGG 16827

Qy       9298 GCCGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCG 9357

Db      16828 GCCGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCG 16887

Qy       9358 GATAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCC 9417

Db      16888 GATAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCC 16947

Qy       9418 ACCCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCG 9477

Db      16948 ACCCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCG 17007

Qy       9478 TGGGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAA 9537

Db      17008 TGGGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAA 17067

Qy       9538 TTTGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAG 9597

Db      17068 TTTGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAG 17127

Qy       9598 CCCGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATC 9657

Db      17128 CCCGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATC 17187

Qy       9658 ACATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGG 9717

Db      17188 ACATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGG 17247

Qy       9718 TTCCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGT 9777

Db      17248 TTCCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGT 17307

Qy       9778 GGCTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTA 9837

Db      17308 GGCTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTA 17367

Qy       9838 GAAGGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAAT 9897

Db      17368 GAAGGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAAT 17427

Qy       9898 CCAGGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGT 9957

Db      17428 CCAGGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGT 17487

Qy       9958 TCTTGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000

Db      17488 TCTTGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 17530
```

RESULT 3
US-10-177-293-505
; Sequence 505, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 505
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-505

Query Match 16.4%; Score 1639.4; DB 15; Length 5632;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 2533

Db          1 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 60

Qy       2534 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 2593

Db         61 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 120

Qy       2594 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 2653

Db        121 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 180

Qy       2654 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 2713
```

```
Db        181 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 240

Qy       2714 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773

Db        241 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300

Qy       2774 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 2833

Db        301 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 360

Qy       2834 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 2893

Db        361 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 420

Qy       2894 TACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 2953

Db        421 TACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 480

Qy       2954 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 3013

Db        481 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 540

Qy       3014 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073

Db        541 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600

Qy       3074 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 3133

Db        601 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 660

Qy       3134 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 3193

Db        661 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 720

Qy       3194 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 3253

Db        721 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 780

Qy       3254 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 3313

Db        781 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 840

Qy       3314 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 3373

Db        841 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 900

Qy       3374 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 3433

Db        901 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 960

Qy       3434 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 3493

Db        961 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 1020

Qy       3494 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 3553

Db       1021 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 1080

Qy       3554 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 3613

Db       1081 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 1140

Qy       3614 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 3673

Db       1141 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 1200

Qy       3674 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 3733

Db       1201 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 1260

Qy       3734 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 3793

Db       1261 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 1320

Qy       3794 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 3853

Db       1321 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 1380

Qy       3854 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 3913

Db       1381 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 1440

Qy       3914 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 3973

Db       1441 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 1500

Qy       3974 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 4033

Db       1501 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 1560

Qy       4034 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 4093

Db       1561 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 1620

Qy       4094 CGAAGGAATCCATCTGGGTAA 4114

Db       1621 CGAAGGAATCCATCTGGATAA 1641
```

```
RESULT 4
US-08-731-499-10
; Sequence 10, Application US/08731499
; Publication No. US20030148270A1
;  GENERAL INFORMATION:
;    APPLICANT:  GRAY, Joe W.
;    APPLICANT:  COLLINS, Colin
;    APPLICANT:  HWANG, Soo-In
;    APPLICANT:  GODFREY, Tony
;    APPLICANT:  KOWBEL, David
;    APPLICANT:  ROMMENS, Johanna
;    TITLE OF INVENTION:  GENES FROM THE 20q13 AMPLICON AND THEIR
;    TITLE OF INVENTION:  USES
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/731,499
;      FILING DATE:  16-OCT-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/680,395
;      FILING DATE:  15-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hunter, Tom
;      REGISTRATION NUMBER:  38,498
;      REFERENCE/DOCKET NUMBER:  23070-068910
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3186 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
```

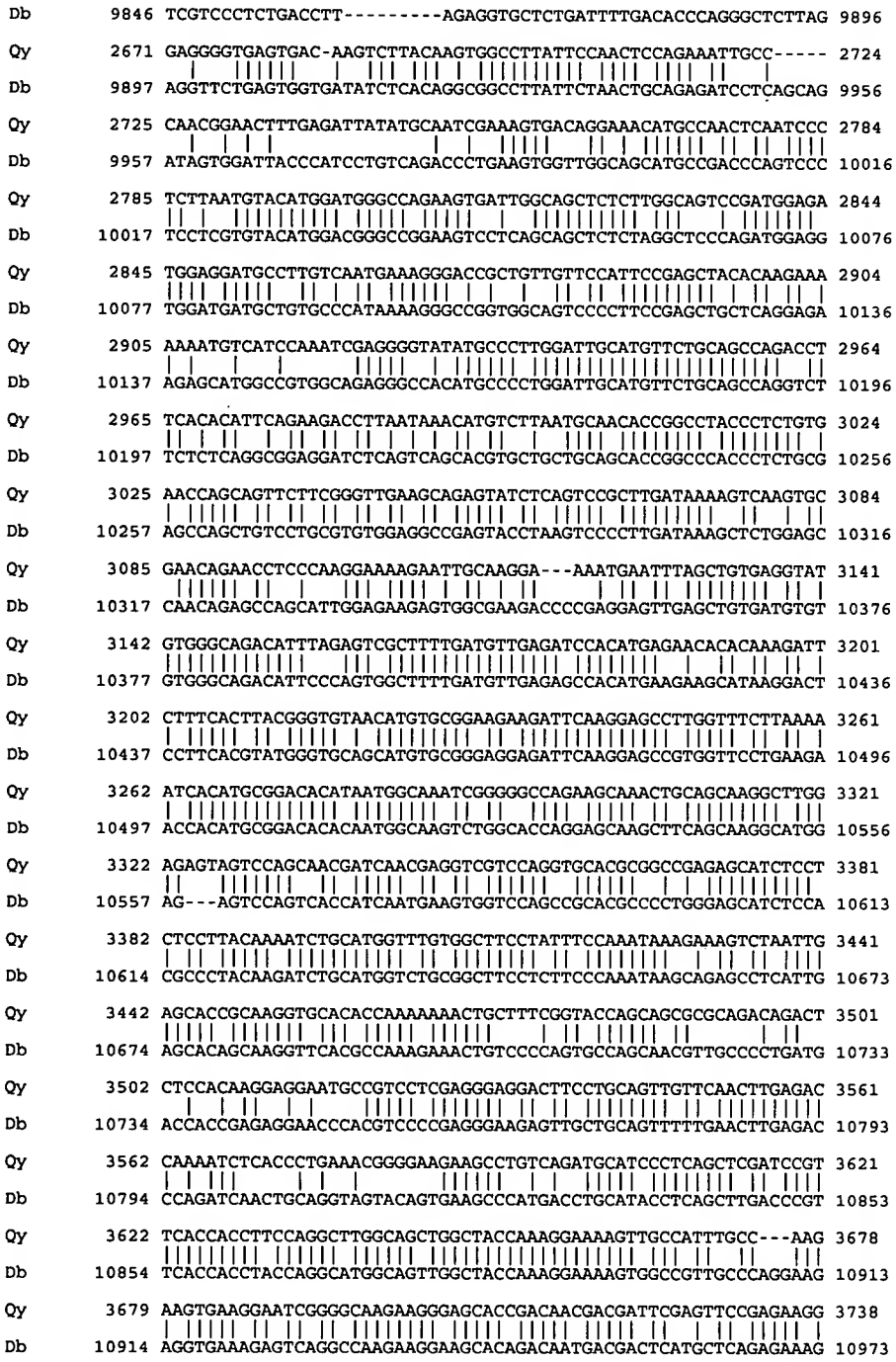
```
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1..3186
;      OTHER INFORMATION:  /note= "ZABC1 Open Reading Frame"
US-08-731-499-10

  Query Match             15.6%;  Score 1559.6;  DB 8;  Length 3186;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1565;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       8286 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACATCTCT 8345
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1482 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACATCTCT 1541

Qy       8346 GAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 8405
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1542 GAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 1601

Qy       8406 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCGCAAAC 8465
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1602 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCGCAAAC 1661

Qy       8466 CAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCTGCAAA 8525
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1662 CAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCTGCAAA 1721

Qy       8526 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTCTCACC 8585
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1722 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTCTCACC 1781

Qy       8586 AGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGATAAAGT 8645
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1782 AGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGATAAAGT 1841

Qy       8646 GAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAC 8705
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1842 GAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAC 1901

Qy       8706 TCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGATGGCAG 8765
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1902 TCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGATGGCAG 1961

Qy       8766 TACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCTGACTG 8825
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1962 TACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCTGACTG 2021

Qy       8826 CAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCT 8885
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2022 CAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCT 2081

Qy       8886 TCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGTCCATT 8945
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2082 TCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGTCCATT 2141

Qy       8946 TTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAGCATAA 9005
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2142 TTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAGCATAA 2201

Qy       9006 ATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGACGTAC 9065
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2202 ATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGACGTAC 2261

Qy       9066 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGTAAACC 9125
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       2262 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCTCTAGTTTCTGTAAACC 2321

Qy       9126 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 9185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2322 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 2381

Qy       9186 GAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 9245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2382 GAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 2441

Qy       9246 CCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCCGCCAC 9305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2442 CCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCCGCCAC 2501

Qy       9306 CAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGATAAGAC 9365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2502 CAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGATAAGAC 2561

Qy       9366 AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACCCTCGG 9425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2562 AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACCCTCGG 2621

Qy       9426 CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGGGCACC 9485
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2622 CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGGGCACC 2681

Qy       9486 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTTGGTGA 9545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2682 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTTGGTGA 2741

Qy       9546 GCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCCGGGGC 9605
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2742 GCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCCGGGGC 2801

Qy       9606 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT 9665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2802 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT 2861

Qy       9666 GTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTCCTGAG 9725
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2862 GTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTCCTGAG 2921

Qy       9726 CTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGCTCCGG 9785
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2922 CTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGCTCCGG 2981

Qy       9786 GCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTAT 9845
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       2982 GCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTCT 3041

Qy       9846 TGCATGAGGGGCGT 9859
             ||   || |   ||
Db       3042 TGGTGGATGTCAGT 3055


RESULT 5
US-10-087-192-1709
; Sequence 1709, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1709
;   LENGTH: 5252
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-087-192-1709

  Query Match             15.6%;  Score 1559.6;  DB 13;  Length 5252;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1565;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;
```

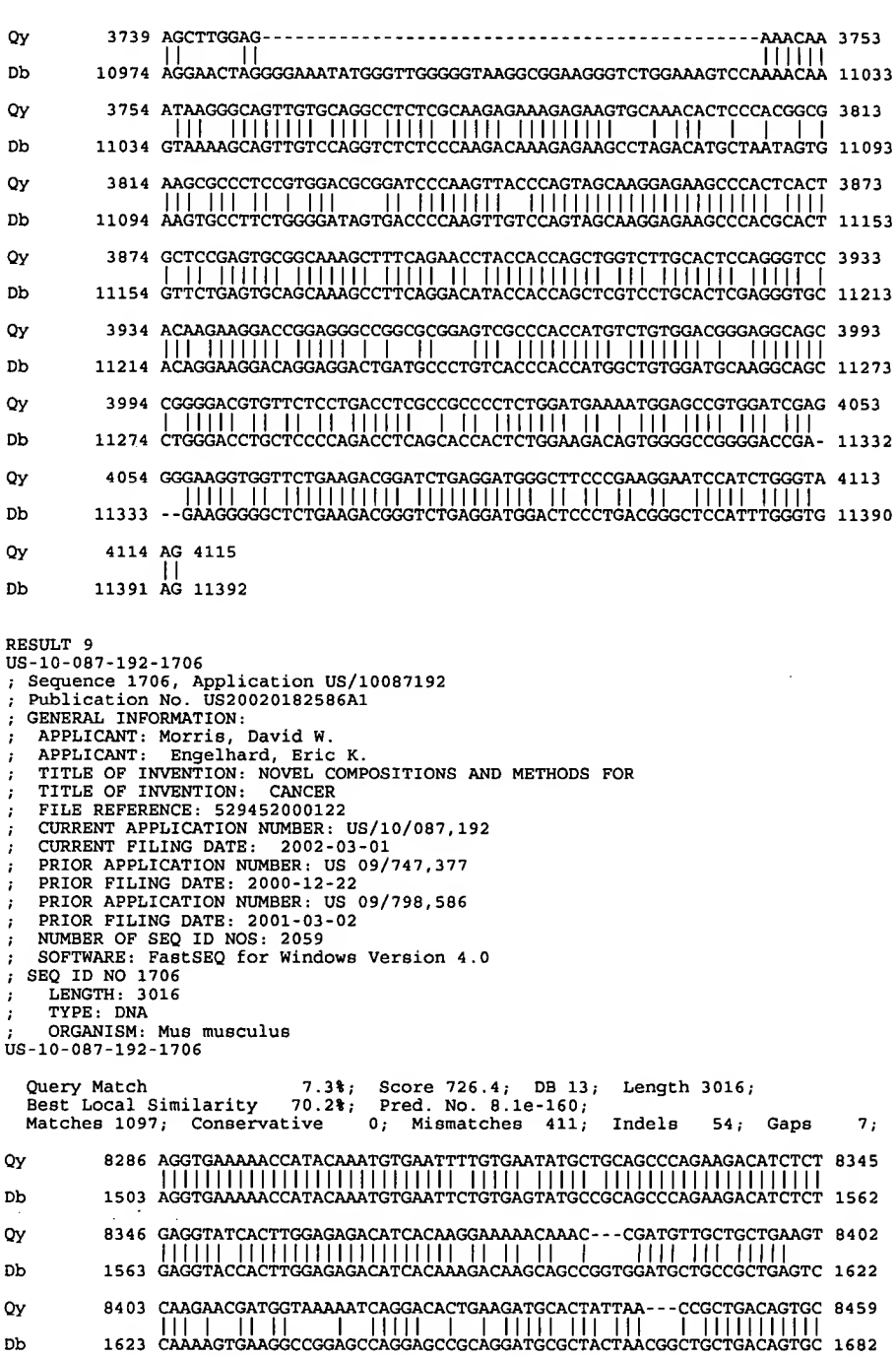
```
Qy       8286 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACATCTCT 8345
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1506 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACATCTCT 1565

Qy       8346 GAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 8405
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1566 GAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 1625

Qy       8406 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCGCAAAC 8465
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1626 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCGCAAAC 1685

Qy       8466 CAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCTGCAAA 8525
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1686 CAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCTGCAAA 1745

Qy       8526 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTCTCACC 8585
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1746 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTCTCACC 1805

Qy       8586 AGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGATAAAGT 8645
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1806 AGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGATAAAGT 1865

Qy       8646 GAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAC 8705
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1866 GAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAC 1925

Qy       8706 TCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGATGGCAG 8765
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1926 TCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGATGGCAG 1985

Qy       8766 TACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCTGACTG 8825
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1986 TACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCTGACTG 2045

Qy       8826 CAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCT 8885
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2046 CAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCT 2105

Qy       8886 TCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGTCCATT 8945
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2106 TCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGTCCATT 2165

Qy       8946 TTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAGCATAA 9005
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2166 TTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAGCATAA 2225

Qy       9006 ATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGACGTAC 9065
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2226 ATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGACGTAC 2285

Qy       9066 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGTAAACC 9125
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       2286 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCTCTAGTTTCTGTAAACC 2345

Qy       9126 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 9185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2346 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 2405

Qy       9186 GAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 9245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2406 GAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 2465

Qy       9246 CCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCCGCCAC 9305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2466 CCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCCGCCAC 2525

Qy       9306 CAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGATAAGAC 9365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2526 CAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGATAAGAC 2585

Qy       9366 AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACCCTCGG 9425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2586 AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACCCTCGG 2645

Qy       9426 CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGGGCACC 9485
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2646 CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGGGCACC 2705

Qy       9486 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTTGGTGA 9545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2706 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTTGGTGA 2765

Qy       9546 GCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCCGGGGC 9605
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2766 GCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCCGGGGC 2825

Qy       9606 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT 9665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2826 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT 2885

Qy       9666 GTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTCCTGAG 9725
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2886 GTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTCCTGAG 2945

Qy       9726 CTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGCTCCGG 9785
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2946 CTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGCTCCGG 3005

Qy       9786 GCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTAT 9845
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       3006 GCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTCT 3065

Qy       9846 TGCATGAGGGGCGT 9859
             ||   || |   ||
Db       3066 TGGTGGATGTCAGT 3079
```

RESULT 6
US-08-731-499-3
; Sequence 3, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; TITLE OF INVENTION: USES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498

```
;      REFERENCE/DOCKET NUMBER:  2307O-068910
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1507 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1..1507
;      OTHER INFORMATION:  /note= "cDNA clone cc49 of 6-7kb
;      OTHER INFORMATION:  transcript with homology to C2H2 zinc
;      OTHER INFORMATION:  finger genes"
US-08-731-499-3

  Query Match             13.9%;  Score 1387.2;  DB 8;  Length 1507;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches 1470;  Conservative    3;  Mismatches   32;  Indels    6;  Gaps    6;

Qy       2424 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGA 2483
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGC 61

Qy       2484 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGT 2543
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGT 121

Qy       2544 TTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGG 2603
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGG 181

Qy       2604 ACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTTGGAACCCAGG 2663
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        182 ACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGG-TTTTGGAACCCAGG 240

Qy       2664 GCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTCCAGAAATTGC 2723
              |||||||| |||||||||||| ||||||| |||||||||| ||||||||||||||||||
Db        241 ACTGTTAGA-GGGTGAGTGACAGGTCTTAC-AGTGGCCTTAATCCAACTCCAGAAATTGC 298

Qy       2724 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 2783
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 358

Qy       2784 CTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAG 2843
             |||||||||||||||| |||||  |||||||||||||||||| ||||||||||||||||||
Db        359 CTCTTAATGTACATGGAT-GGCCAAGAGTGATTGGCAGCTCTCTTGCCAGTCCGATGGAG 417

Qy       2844 ATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGAA 2903
             ||||| |||||||||||||||||||| || |||||||    | |          ||||||
Db        418 ATGGA-GATGCCTTGTCAATGAAAGGGCCCNCTGTTGTCAATTCCGAGCTACACAAAGAA 476

Qy       2904 AAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 2962
             ||||||||| |||| |||||||||| ||||||||||||||||||||||||||||||||||
Db        477 AAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 536

Qy       2963 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 3022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 596

Qy       3023 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 3082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 656

Qy       3083 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 3142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 716

Qy       3143 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 3202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        717 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 776

Qy       3203 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 3262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 836

Qy       3263 TCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 3322
             |||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCACATGCGGACRCATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 896

Qy       3323 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGGAGAGCATCTCCTC 3382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGGAGAGCATCTCCTC 956

Qy       3383 TCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 3442
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 TCCTTGCAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 1016

Qy       3443 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTC 3502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTC 1076

Qy       3503 TCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGACC 3562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 TCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGACC 1136

Qy       3563 AAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGTT 3622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1137 AAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGTT 1196

Qy       3623 CACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTGCCAAGAAGT 3682
             |||||||||||||||||||||:||||||||||||||||:||||||||||||||||||||||
Db       1197 CACCACCTTCCAGGCTTGGCAKCTGGCTACCAAAGGAAWAGTTGCCATTTGCCAAGAAGT 1256

Qy       3683 GAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGAGAAGGAGCT 3742
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1257 GAAGGAATTGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGAGAAGGAGCT 1316

Qy       3743 TGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACA 3802
             |||||||||||||||  |  ||||||||||||||||||||||||||||||||||||||||
Db       1317 TGGAGAAACAAATAAGAACCATTGTGCAGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACA 1376

Qy       3803 CTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAA 3862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1377 CTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAA 1436

Qy       3863 GCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCTGGTCTTGCA 3922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1437 GCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCTGGTCTTGCA 1496

Qy       3923 CTCCAGGGTCC 3933
             ||||||||||| 
Db       1497 CTCCAGGGTCC 1507


RESULT 7
US-10-684-422-156/c
; Sequence 156, Application US/10684422
; Publication No. US20040229233A1
; GENERAL INFORMATION:
;  APPLICANT: ABURATANI, Hiroyuki
;  APPLICANT:  YAMAMOTO, Shogo
;  TITLE OF INVENTION: Human housekeeping genes and human tissue-specific genes
;  FILE REFERENCE: 113991
;  CURRENT APPLICATION NUMBER: US/10/684,422
;  CURRENT FILING DATE:  2002-10-15
;  PRIOR APPLICATION NUMBER: US 60/418,614
;  PRIOR FILING DATE: 2002-10-16
;  NUMBER OF SEQ ID NOS: 332
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 156
;   LENGTH: 124990
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
US-10-684-422-156

  Query Match             10.1%;  Score 1006.8;  DB 20;  Length 124990;
  Best Local Similarity   95.5%;  Pred. No. 6.3e-225;
  Matches 1083;  Conservative    0;  Mismatches   37;  Indels   14;  Gaps    4;

Qy          1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGG------AGACTTGCTCTGTTGCCCA 54
              |||||||||||||| |||||| ||    ||| ||||      || || ||||||||||||
Db       1128 CCATCATATTTCTTTTTTTTTGGGCAAGAGGGGGGAGTCAAAGTCTCGCTCTGTTGCCCA 1069

Qy         55 GGCTGGA--CCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTG 112
              |||||||   |||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1068 GGCTGGAATGCAGTGGCGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTG 1009

Qy        113 ATTCCCAAATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTT 172
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1008 ATTCCCAAGTAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTT 949

Qy        173 TAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTG 232
              |||||||||||||||||||||||||| ||||||||||||||||||| ||||| |||
Db        948 TAGCAGATAAGGGGTTTCACCATGTTGGTCAGGCTGGTCTCCAACTCCTGACCTCAAGTG 889

Qy        233 ATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTC 292
              |||||||||||||||||||||||||||||||| |||||||||||||||| |||| |
Db        888 ATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTACAGGTGTGAGCCACTATGCCCGGC 829

Qy        293 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGT 352
              |||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db        828 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTGGAAAGGATGT 769

Qy        353 CAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAAT 412
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        768 CAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACATATTAGTCAAGTTTCTAAAAT 709

Qy        413 ACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTT 472
              |||||||||||||||||||| |||||||||||| |     ||||| |||||||||
Db        708 ACAGTAATAATGTTGAGAGCTGACGTAGGGACTAACTTAG-----TTTTTGTTTTTTTTT 654

Qy        473 TTTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAACCAAAA 532
              |||||||||||||  |||||||||||||||||||||||||| |||||||| ||||||
Db        653 TTTCAAATTCTCACCTAACTTTGATTTTGCTAAATAAGGACATT-AAAAAAAACCCAAAA 595

Qy        533 AACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCAT 592
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        594 AACTCCACTATTGCATATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCAT 535

Qy        593 CTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGT 652
              ||  |||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        534 CTGCAAGTAGGAAGGGCCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGT 475

Qy        653 TGTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCC 712
              |||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        474 TGTCAGCATTTGTTAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCC 415

Qy        713 TGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATG 772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATG 355

Qy        773 TATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAAT 832
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        354 TATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAGT 295

Qy        833 GTGTAGCTCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTA 892
              |||||| ||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        294 GTGTAGATCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTAGAGTTTA 235

Qy        893 ACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGT 952
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 ACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGT 175

Qy        953 TGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGC 1012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        174 TGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGC 115

Qy       1013 CCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCT 1072
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        114 CCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAGAGACTTCAGGACATTCTTTCCT 55

Qy       1073 TTACCCTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATC 1126
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db          54 TTACCCTAATCCTTGATCTTCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATC 1


RESULT 8
US-10-087-192-1705
; Sequence 1705, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1705
;   LENGTH: 26345
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-087-192-1705

  Query Match             8.2%;  Score 820;  DB 13;  Length 26345;
  Best Local Similarity   68.1%;  Pred. No. 2.6e-181;
  Matches 1391;  Conservative    0;  Mismatches  550;  Indels  101;  Gaps   14;

Qy       2132 AACTTGACATTTTAGATTTAAGTCGGTAAAGCACTGATTTAAACTGGATTTTAACTGGAT 2191
              || ||     |||| |||| |||  | | ||| |||  ||||       |     | ||
Db       9394 AAATTTGGCATTTATATTTCAGTGAGCAGAGCTCTGCCTTAAGTCTCCTGACTGGTTGAA 9453

Qy       2192 GAAATTCTGATTTAATAAGTGTACTGACTGGATAAAATGCCAATGATTTAATTAACAAGC 2251
                | || | ||| ||||||||   |||||  || |||| |||||||||||||||| |
Db       9454 ATTCTACTCACTTAGTAAGTGTATCCACTGGGAAATATGCTAATGATTTAATTAACAAAC 9513

Qy       2252 ACGTTTAACAGGATGCCCTATATATTAGTTAAAAGTGAAGCAATTGAATTAGGTACCTTC 2311
              | ||||| |||| |||  |||  |||||||||| |||| ||| |||| |   || ||||
Db       9514 ATGTTTAGCAGGCTGCATTATGGATTAGTTAAAGGTGAGGCAGTTGACTCTAGTCCCTTT 9573

Qy       2312 TCTGCTGCGTGGAAAAGACCGTATGACTCACCCACACCAGCCTTCTCTTCGCTCTGAGTG 2371
              |       ||||| |||| | | |||||||     |     ||| || |    ||   |
Db       9574 T-------GTGGAGAAGATCATGTGACTCATAGGCCAAGCCCTGTTCCTGAGCCTTCCTC 9626

Qy       2372 TAGCTAACCGTTTCTGTTTTTTTTCCTCTAGGGTTTGGAAATCCCTTGTCTCCAGGTTGC 2431
                 ||      |||||||||| | |  ||  |||||||||||||||||| ||  ||||
Db       9627 ACTGAGAC-----TTGTTTTTTTCCTTTCAGATTTTGGAAATCCCTTGTCTTCACATTGC 9681

Qy       2432 TGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGAACTAATG- 2490
              |||||  || |||||||||| |||  ||||||  || ||||||    | ||| ||  ||
Db       9682 CGGGATCAACGTCTTGCTCAACTGAGCCACTCA-CCAGTGGAGATGGACAGAGCTGTTGC 9740

Qy       2491 CTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGTTTGTGGA 2550
              | | |||||||   || || ||| ||    || | ||| | | || ||||||||||||| |
Db       9741 CCTCGTGCTGACGCACATCTGACTCACAGCCTGGGGCAGCTTGGATGCCTTGTTTGTGCA 9800

Qy       2551 AAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGGACTATAC 2610
              |||| |||    | ||||||||||||  |  || | |           ||| |||||
Db       9801 AAGAGCCA---TCTCCATCACTGAGCTCTCCCAAGCTCC-----------GGAATATAC 9845

Qy       2611 ACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTTGGAACCCAGGGCTGTTA 2670
              | | | | ||| |||         | |  ||||||| ||||   | ||  ||    |
```





RESULT 9
US-10-087-192-1706
; Sequence 1706, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1706
; LENGTH: 3016
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-1706

Query Match 7.3%; Score 726.4; DB 13; Length 3016;
Best Local Similarity 70.2%; Pred. No. 8.1e-160;
Matches 1097; Conservative 0; Mismatches 411; Indels 54; Gaps 7;

```
Qy       8460 GCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACC 8519
             ||| ||||||||||| || |||||| ||||||||||||||||||||||  || || |||||
Db       1683 GCAGACCAAAAATTTAAAGAGATTTCTTGATGGTGCCAAAGATGTTAAGGGAAGCCCACC 1742

Qy       8520 TGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCT 8579
             ||| ||||||||||||||||||||||||| || |||| |||||             ||
Db       1743 TGCCAAGCAGCTTAAGGAGATGCCTTCTGTCTTCCAGAGTGTTC-----------TCTC 1790

Qy       8580 CTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGA 8639
                || ||     ||| |||||||||||||||||||||| ||||||||||   ||||||||
Db       1791 ACCAGCACACAGCAACGATACTCAGGATTTCCATAAACATGCAGCTGAT---AGTGCTGA 1847

Qy       8640 TAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGT 8699
              |||| ||  || | |||| ||||| |||| |||||| ||    | |||||    |||||
Db       1848 GAAAGCGAGGAAGAGCCCTGCCCCTACTTATCTGGACATGCAGAGAAAGA---AAGCAGG 1904

Qy       8700 TGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGA 8759
              ||  ||||||| |  | ||   |||| |||    ||| || ||   |    | | |||
Db       1905 GGAGCCTCAGGCCAGCAGCCCTGTCTGCAGACTAGAGGGGGTTGGGTCCTTAGCACGGGA 1964

Qy       8760 TGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGC 8819
              |   |                          ||   | || | | |||    |  ||
Db       1965 AGCTGG--------------------------CCATAGGGAGAAGATGGATCAGGATGC 1997

Qy       8820 TGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGG 8879
              ||||| |||| | | |||  |||  || || || || |  ||||| ||| ||||| | ||
Db       1998 TGACTACAGACATAAGCCCGGTGCTGACTGCCAGGACAGGCCTTTGAATCTATCCCTTGG 2057

Qy       8880 GGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTG 8939
              | | || |||   || || ||||| |||||||| ||  || |||| || || ||  ||||
Db       2058 GCCGCTCCACGCCTGTCCTGCAATCTCTTTGAGCAAGTGTCTGATCCCCAGCATTGCCTG 2117

Qy       8940 TCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGA 8999
              || |||||||| |||||||| |||||||| |||||  | |||||||||||||||||| ||
Db       2118 CCCCTTTTGTACTTTCAAGACCTTTTATCCGGAAGTCCTTATGATGCACCAGAGACTTGA 2177

Qy       9000 GCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCG 9059
              ||| |  ||||| ||||||    || || ||| |  | | ||||||  |||| || |   |
Db       2178 GCACAGGTACAACCCTGACCCGCACAAGAACGGCAGCAGCAAGTCTGTGCTGAGGAACAG 2237

Qy       9060 ACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTG 9119
              |||||||| ||||| || || |||||||||||||||||| ||||||  |  || |  |
Db       2238 GCGTACCGGGTGCCCTCCGGCTTTGCTGGGAAAAGATGTCCCTCCCTTGTCTGGCCTGCA 2297

Qy       9120 TAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGG 9179
              || |||||| ||||| |||| ||| |  | || || || |||||||  || | |||||
Db       2298 CAAGCCCAAGGCCAAGACTGCCTTCTCACCACACTCGAAGTCCCTGCACTCAGAGAAGGC 2357

Qy       9180 GAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCAC 9239
                |||| |  |  | |||||| |||| || || | |||||||||   ||||  |||||
Db       2358 TCGGCAGGGGGCCTCGGGGCCAAGCAAAGCACCCCAGACTTCAGGACCAGACAACAGCAC 2417

Qy       9240 TTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGC 9299
              |||||||||||||||||||||||| |||||  |||| |  ||||  |||| |     |
Db       2418 TTTAGCCCCAAGTAACCTGAAGTCACACAGGTCACAACCCAATGCTGGGGGCACCAG--- 2474

Qy       9300 CGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGA 9359
              |||||||||| ||||| || ||| |||||| |||    |  ||  |  ||||   |||
Db       2475 TGCCACCAGGCAGCAGCAGTCAGAGTTGTTTCCCAAAGGTGGCGTCCCTGCTGCTATGGA 2534

Qy       9360 TAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCAC 9419
              ||||   || ||||| ||| |||| |||||  | || ||||    | || |||||| || |
Db       2535 TAAGGTGAAGAGACCTGAGCCAAAACTGAAGTCCCTACCAGCCTCCCCGTCTCAGTCCCC 2594

Qy       9420 CCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTG 9479
              |||| |||| | ||| | ||| |||||  | || || ||||  |||   ||| |||| ||
Db       2595 CCTCAGCAGTAATAATAGCAACGGTTCTGTTGAGTATCCCGTGAAGGTTGACGGCCCATG 2654

Qy       9480 GGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATT 9539
              |||||  |  || |||||||| | |||  |||||| | |  ||| | | |||||||| |
Db       2655 GGCACAGCAAGGGAGAGACTACTACTGCCATCGGAATTCTGGCAGTGCCGCAGCAGAGTA 2714

Qy       9540 TGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCC 9599
              ||||||| | ||| |||||||| |||||||| | |||  |||| |||   || ||  |
Db       2715 CAGTGAGCCACATCCCAAAAGACTCAAGTCCAGTGCGGTGTCCCTGGACACAGAGCATGC 2774

Qy       9600 CGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCAC 9659
              ||| |||||  |||||| |||| ||| || |||||||||||| |||||||  |||||||
Db       2775 AGGGACCAATGGCAGAAGGGGCTTTGAGCTCCCCAAGTACCATGTGGTCAGGAGCATCAC 2834

Qy       9660 ATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTT 9719
              ||  || |||| |  || ||||||   ||  |||  | |||||| | |||| |  | ||
Db       2835 CTCCTTGCTACCACCAGAGTGTGTGCGCCCACCGCCTGTGCTGCCCCACAAAGCCCGTTT 2894

Qy       9720 CCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGG 9779
              ||||||| |  | ||||| || || || | |||| || |||| ||||||||||  |||
Db       2895 CCTGAGCCCTGGGGAGGTGGAGTCACCCAGTGTGTTGGCTGTGCAGAAGCCCTACAGTGC 2954

Qy       9780 CTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGA 9839
              ||| || || || || || |||| || | ||     ||| |  |||| |  |  | ||
Db       2955 CTCTGGACCCCTGTATACCTGTGGACCCGTGGGACACGCAGGAGGCAGCCCAGCCCTTGA 3014

Qy       9840 AG 9841
              ||
Db       3015 AG 3016
```

RESULT 10
US-10-027-632-104193
; Sequence 104193, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 104193
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-104193

Query Match 6.2%; Score 617.2; DB 13; Length 707;
Best Local Similarity 94.4%; Pred. No. 1.8e-134;
Matches 674; Conservative 1; Mismatches 29; Indels 10; Gaps 3;

```
Qy         39 CTTGCTCTGTTGCCCAGGCTGGA--CCAGTGGTGCGATCTTGGCTCACTGCAACCTCCAC 96
              || || |||||||||||||||||   ||||||  |||||||||||||||| ||||||||
Db          2 CTCGCCCTGTTGCCCAGGCTGGAATGCAGTGGCACGATCTTGGCTCACTGCTACCTCCAC 61

Qy         97 CTCCTGGGTTCAAGTGATTCCCAAATAGCTGGGATTACAGGTGTGTATTACCATGCCCAG 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TTCCTGGGTTCAAGTGATTCCCAAGTAGCTGGGATTACAGGTGTGTATTACCATGCCCAG 121

Qy        157 CTAATTTTTGTATTTTTAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAA 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CTAATTTTTGTATTTTTAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAA 181
```

```
Qy        217 CTCCTGGCCTCATGTGATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTG 276
              |||||||||||| |||||||||||||||||||| |||||||||||||||||| ||| |||
Db        182 CTCCTGGCCTCAAGTGATCCACCCACTTCGGCTACCCAAAGCATTGGGAGTACAGGGGTG 241

Qy        277 AGCCACTATACCCGTCCTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTC 336
              ||||||||| |||  ||||||||||||||||||||||||||||||||||||||||||||
Db        242 AGCCACTATGCCCAGCCTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTC 301

Qy        337 TTTTTCTAAAGGATGTCAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTA 396
              |||||  |||||||||||||||||| |||||||||||||||||||||||||| |||||
Db        302 TTTTTGGAAAGGATGTCAGTAGAGAAATGGAGTTCCCCAAAATTACAGTTTCACATATTA 361

Qy        397 GTCAAGTTTCTAAAATACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTGGTTTT 456
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        362 GTAAAGTTTCTAAAATACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTAG---- 417

Qy        457 TTTTTTTTTTTTTTTTTTTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGACATT 516
                 |||||||||||||||||||| ||||  |||||||||||||||||||||||||||||
Db        418 ---TTTTTTTTTTTTTTTTCAAATTGTCACCTAACTTTGATTTTGCTAAATAAGGACATT 474

Qy        517 AAAAAAAAAACCAAAAAACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAAAAT 576
               |||| |||||||||||| |||||||||| ||||||||||||||||||||||||||||||
Db        475 -AAAAGAAAACCAAAAAAGTCCACTATTGGCTATTGCCACTATTTGATTTTTTAAAAAAT 533

Qy        577 AAGCGTATTTTAGCATCTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGG 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 AAGCGTATTTTAGCATCTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGG 593

Qy        637 CCAGGGAAAACAGCGTTGTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATT 696
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        594 CCAGGGAAAACAGCGTTGTCAGCATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATT 653

Qy        697 CAGTTAAAACCTTGCCTGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGC 750
              ||||||||||:|||||||||||||||||||| ||||||||||||||||||||
Db        654 CAGTTAAAACSTTGCCTGGGACGCTAGCATTCGGTAAATACTTGTTGAATAAGC 707
```

RESULT 11
US-10-027-632-325113
; Sequence 325113, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 325113
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-325113

Query Match 6.2%; Score 617.2; DB 13; Length 707;
Best Local Similarity 94.4%; Pred. No. 1.8e-134;

Matches 674; Conservative 1; Mismatches 29; Indels 10; Gaps 3;

```
Qy         39 CTTGCTCTGTTGCCCAGGCTGGA--CCAGTGGTGCGATCTTGGCTCACTGCAACCTCCAC 96
              || || |||||||||||||||||   |||||| |||||||||||||||| ||||||||
Db          2 CTCGCCCTGTTGCCCAGGCTGGAATGCAGTGGCACGATCTTGGCTCACTGCTACCTCCAC 61

Qy         97 CTCCTGGGTTCAAGTGATTCCCAAATAGCTGGGATTACAGGTGTGTATTACCATGCCCAG 156
              ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db         62 TTCCTGGGTTCAAGTGATTCCCAAGTAGCTGGGATTACAGGTGTGTATTACCATGCCCAG 121

Qy        157 CTAATTTTTGTATTTTTAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAA 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CTAATTTTTGTATTTTTAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAA 181

Qy        217 CTCCTGGCCTCATGTGATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTG 276
              |||||||||||| |||||||||||||||||||| |||||||||||||||||| ||| |||
Db        182 CTCCTGGCCTCAAGTGATCCACCCACTTCGGCTACCCAAAGCATTGGGAGTACAGGGGTG 241

Qy        277 AGCCACTATACCCGTCCTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTC 336
              ||||||||| |||  ||||||||||||||||||||||||||||||||||||||||||||
Db        242 AGCCACTATGCCCAGCCTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTC 301

Qy        337 TTTTTCTAAAGGATGTCAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTA 396
              |||||  |||||||||||||||||| |||||||||||||||||||||||||| |||||
Db        302 TTTTTGGAAAGGATGTCAGTAGAGAAATGGAGTTCCCCAAAATTACAGTTTCACATATTA 361

Qy        397 GTCAAGTTTCTAAAATACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTGGTTTT 456
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        362 GTAAAGTTTCTAAAATACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTAG---- 417

Qy        457 TTTTTTTTTTTTTTTTTTTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGACATT 516
                 |||||||||||||||||||| ||||  |||||||||||||||||||||||||||||
Db        418 ---TTTTTTTTTTTTTTTTCAAATTGTCACCTAACTTTGATTTTGCTAAATAAGGACATT 474

Qy        517 AAAAAAAAAACCAAAAAACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAAAAT 576
               |||| |||||||||||| |||||||||| ||||||||||||||||||||||||||||||
Db        475 -AAAAGAAAACCAAAAAAGTCCACTATTGGCTATTGCCACTATTTGATTTTTTAAAAAAT 533

Qy        577 AAGCGTATTTTAGCATCTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGG 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 AAGCGTATTTTAGCATCTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGG 593

Qy        637 CCAGGGAAAACAGCGTTGTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATT 696
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        594 CCAGGGAAAACAGCGTTGTCAGCATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATT 653

Qy        697 CAGTTAAAACCTTGCCTGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGC 750
              ||||||||||:|||||||||||||||||||| ||||||||||||||||||||
Db        654 CAGTTAAAACSTTGCCTGGGACGCTAGCATTCGGTAAATACTTGTTGAATAAGC 707
```

RESULT 12
US-10-027-632-104193
; Sequence 104193, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358

```
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 104193
;   LENGTH: 707
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-104193

  Query Match             6.2%;  Score 617.2;  DB 17;  Length 707;
  Best Local Similarity   94.4%;  Pred. No. 1.8e-134;
  Matches  674;  Conservative    1;  Mismatches   29;  Indels   10;  Gaps    3;

Qy         39 CTTGCTCTGTTGCCCAGGCTGGA--CCAGTGGTGCGATCTTGGCTCACTGCAACCTCCAC 96
              || || ||||||||||||||||    ||||||  |||||||||||||||| ||||||||
Db          2 CTCGCCCTGTTGCCCAGGCTGGAATGCAGTGGCACGATCTTGGCTCACTGCTACCTCCAC 61

Qy         97 CTCCTGGGTTCAAGTGATTCCCAAATAGCTGGGATTACAGGTGTGTATTACCATGCCCAG 156
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         62 TTCCTGGGTTCAAGTGATTCCCAAGTAGCTGGGATTACAGGTGTGTATTACCATGCCCAG 121

Qy        157 CTAATTTTTGTATTTTTAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAA 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CTAATTTTTGTATTTTTAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAA 181

Qy        217 CTCCTGGCCTCATGTGATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTG 276
              |||||||||||| |||||||||||||||||| |||||||||||||||| ||| |||
Db        182 CTCCTGGCCTCAAGTGATCCACCCACTTCGGCTACCCAAAGCATTGGGAGTACAGGGGTG 241

Qy        277 AGCCACTATACCCGTCCTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTC 336
              ||||||||| |||  ||||||||||||||||||||||||||||||||||||||||||||
Db        242 AGCCACTATGCCCAGCCTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTC 301

Qy        337 TTTTTCTAAAGGATGTCAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTA 396
              |||||  |||||||||||||||||| ||||||||||||||||||||||||||||| |||||
Db        302 TTTTTGGAAAGGATGTCAGTAGAGAAATGGAGTTCCCCAAAATTACAGTTTCACATATTA 361

Qy        397 GTCAAGTTTCTAAAATACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTGGTTTT 456
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        362 GTAAAGTTTCTAAAATACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTAG---- 417

Qy        457 TTTTTTTTTTTTTTTTTTTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGACATT 516
              |||||||||||||||||||| ||||  ||||||||||||||||||||||||||||||||
Db        418 ---TTTTTTTTTTTTTTTTCAAATTGTCACCTAACTTTGATTTTGCTAAATAAGGACATT 474

Qy        517 AAAAAAAAAACCAAAAAACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAAAAT 576
              |||| |||||||||||| |||||||||| ||||||||||||||||||||||||||||||
Db        475 -AAAAGAAAACCAAAAAAGTCCACTATTGGCTATTGCCACTATTTGATTTTTTAAAAAAT 533

Qy        577 AAGCGTATTTTAGCATCTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGG 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 AAGCGTATTTTAGCATCTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGG 593

Qy        637 CCAGGGAAAACAGCGTTGTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATT 696
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        594 CCAGGGAAAACAGCGTTGTCAGCATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATT 653

Qy        697 CAGTTAAAACCTTGCCTGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGC 750
              ||||||||||:|||||||||||||||||||| ||||||||||||||||||||
Db        654 CAGTTAAAACSTTGCCTGGGACGCTAGCATTCGGTAAATACTTGTTGAATAAGC 707


RESULT 13
US-10-027-632-325113
; Sequence 325113, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 325113
;   LENGTH: 707
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-325113

  Query Match             6.2%;  Score 617.2;  DB 17;  Length 707;
  Best Local Similarity   94.4%;  Pred. No. 1.8e-134;
  Matches  674;  Conservative    1;  Mismatches   29;  Indels   10;  Gaps    3;

Qy         39 CTTGCTCTGTTGCCCAGGCTGGA--CCAGTGGTGCGATCTTGGCTCACTGCAACCTCCAC 96
              || || ||||||||||||||||    ||||||  |||||||||||||||| ||||||||
Db          2 CTCGCCCTGTTGCCCAGGCTGGAATGCAGTGGCACGATCTTGGCTCACTGCTACCTCCAC 61

Qy         97 CTCCTGGGTTCAAGTGATTCCCAAATAGCTGGGATTACAGGTGTGTATTACCATGCCCAG 156
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         62 TTCCTGGGTTCAAGTGATTCCCAAGTAGCTGGGATTACAGGTGTGTATTACCATGCCCAG 121

Qy        157 CTAATTTTTGTATTTTTAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAA 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CTAATTTTTGTATTTTTAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAA 181

Qy        217 CTCCTGGCCTCATGTGATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTG 276
              |||||||||||| |||||||||||||||||| |||||||||||||||| ||| |||
Db        182 CTCCTGGCCTCAAGTGATCCACCCACTTCGGCTACCCAAAGCATTGGGAGTACAGGGGTG 241

Qy        277 AGCCACTATACCCGTCCTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTC 336
              ||||||||| |||  ||||||||||||||||||||||||||||||||||||||||||||
Db        242 AGCCACTATGCCCAGCCTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTC 301

Qy        337 TTTTTCTAAAGGATGTCAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTA 396
              |||||  |||||||||||||||||| ||||||||||||||||||||||||||||| |||||
Db        302 TTTTTGGAAAGGATGTCAGTAGAGAAATGGAGTTCCCCAAAATTACAGTTTCACATATTA 361

Qy        397 GTCAAGTTTCTAAAATACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTGGTTTT 456
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        362 GTAAAGTTTCTAAAATACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTAG---- 417

Qy        457 TTTTTTTTTTTTTTTTTTTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGACATT 516
              |||||||||||||||||||| ||||  ||||||||||||||||||||||||||||||||
Db        418 ---TTTTTTTTTTTTTTTTCAAATTGTCACCTAACTTTGATTTTGCTAAATAAGGACATT 474

Qy        517 AAAAAAAAAACCAAAAAACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAAAAT 576
              |||| |||||||||||| |||||||||| ||||||||||||||||||||||||||||||
Db        475 -AAAAGAAAACCAAAAAAGTCCACTATTGGCTATTGCCACTATTTGATTTTTTAAAAAAT 533

Qy        577 AAGCGTATTTTAGCATCTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGG 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 AAGCGTATTTTAGCATCTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGG 593

Qy        637 CCAGGGAAAACAGCGTTGTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATT 696
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        594 CCAGGGAAAACAGCGTTGTCAGCATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATT 653

Qy        697 CAGTTAAAACCTTGCCTGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGC 750
```

```
Db        654 CAGTTAAAACSTTGCCTGGGACGCTAGCATTCGGTAAATACTTGTTGAATAAGC 707


RESULT 14
US-10-029-386-9711/c
; Sequence 9711, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 9711
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO CHR20.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94
;   OTHER INFORMATION: SWISSPROT HIT: O75362, EVALUE 2.00e-67
;   OTHER INFORMATION: NT HIT: gi11421959, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: BF237807.1, EVALUE 0.00e+00
US-10-029-386-9711

  Query Match             5.3%;  Score 530;  DB 16;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-114;
  Matches  530;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3712 ACAACGACGATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAG 3771
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        530 ACAACGACGATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAG 471

Qy       3772 GCCTCTCGCAAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACG 3831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        470 GCCTCTCGCAAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACG 411

Qy       3832 CGGATCCCAAGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAG 3891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        410 CGGATCCCAAGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAG 351

Qy       3892 CTTTCAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGG 3951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 CTTTCAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGG 291

Qy       3952 CCGGCGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTG 4011
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 CCGGCGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTG 231

Qy       4012 ACCTCGCCGCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAG 4071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 ACCTCGCCGCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAG 171

Qy       4072 ACGGATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTC 4131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 ACGGATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTC 111

Qy       4132 CCGTGCTGTTCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCC 4191
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 CCGTGCTGTTCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCC 51

Qy       4192 AGACAACGCTGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGG 4241
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         50 AGACAACGCTGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGG 1


RESULT 15
US-10-040-739-520
; Sequence 520, Application US/10040739
; Publication No. US20020173635A1
;  GENERAL INFORMATION:
;       APPLICANT: Jacobs, Kenneth
;                  McCoy, John
;                  LaVallie, Edward
;                  Racie, Lisa
;                  Merberg, David
;                  Treacy, Maurice
;                  Spaulding, Vikki
;       TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
;       NUMBER OF SEQUENCES: 1519
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genetics Institute, Inc.
;            STREET: 87 CambridgePark Drive
;            CITY: Cambridge
;            STATE: Massachusetts
;            COUNTRY: U.S.A
;            ZIP: 02140
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy Disk
;            COMPUTER: IBM PC Compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/040,739
;            FILING DATE: 07-Jan-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/036,520
;            FILING DATE: 03-JUN-1998
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Brown, Scott A.
;            REGISTRATION NUMBER: 32,724
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617) 498-8224
;            TELEFAX: (617) 876-5851
;  INFORMATION FOR SEQ ID NO: 520:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 469 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 520:
US-10-040-739-520

  Query Match             3.4%;  Score 335.8;  DB 13;  Length 469;
  Best Local Similarity   99.4%;  Pred. No. 3.1e-68;
  Matches  337;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3776 CTCGCAAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGA 3835
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          19 CTAGCAAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGA 78

Qy       3836 TCCCAAGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTT 3895
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          79 TCCCAAGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTT 138

Qy       3896 CAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGG 3955
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         139 CAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGG 198

Qy       3956 CGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCT 4015
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         199 CGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCT 258
```

```
Qy       4016 CGCCGCCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGG 4075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CGCCGCCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGG 318

Qy       4076 ATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGGTAA 4114
              |||||||||||||||||||||||||||||||||||| |||
Db        319 ATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGATAA 357


Search completed: August  1, 2005, 03:00:17
Job time : 3882 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 12:46:07 ; Search time 1023 Seconds
(without alignments)
15994.885 Million cell updates/sec

Title: US-08-731-499-9_COPY_1_10000
Perfect score: 10000
Sequence: 1 CCATCATATTTCTTATTTTT.....ATTCTTAGATACGGCAGTGG 10000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	10000	100.0	10365	4	US-08-892-695-9	Sequence 9, Appli
	2	9509.2	95.1	20022	4	US-09-949-016-12604	Sequence 12604, A
	3	9509.2	95.1	20023	4	US-09-949-016-16004	Sequence 16004, A
	4	1639.4	16.4	5632	3	US-09-560-594-3	Sequence 3, Appli
	5	1639.4	16.4	5632	4	US-09-949-016-862	Sequence 862, App
	6	1639.4	16.4	5632	4	US-09-949-016-4262	Sequence 4262, Ap
	7	1559.6	15.6	3186	4	US-08-892-695-10	Sequence 10, Appl
	8	1387.2	13.9	1507	2	US-08-680-395-3	Sequence 3, Appli
	9	1387.2	13.9	1507	4	US-08-892-695-3	Sequence 3, Appli
c	10	600.6	6.0	601	4	US-09-949-016-34699	Sequence 34699, A
c	11	600.6	6.0	601	4	US-09-949-016-34700	Sequence 34700, A
c	12	600.6	6.0	601	4	US-09-949-016-34701	Sequence 34701, A
c	13	600.6	6.0	601	4	US-09-949-016-34702	Sequence 34702, A
c	14	600.6	6.0	601	4	US-09-949-016-151879	Sequence 151879,
c	15	600.6	6.0	601	4	US-09-949-016-151880	Sequence 151880,
c	16	600.6	6.0	601	4	US-09-949-016-151881	Sequence 151881,
c	17	600.6	6.0	601	4	US-09-949-016-151882	Sequence 151882,
c	18	230.2	2.3	99580	4	US-09-949-016-17411	Sequence 17411, A
	19	228.8	2.3	119981	4	US-09-949-016-11844	Sequence 11844, A
	20	228.8	2.3	119982	4	US-09-949-016-13606	Sequence 13606, A
c	21	228.6	2.3	46841	4	US-09-949-016-13466	Sequence 13466, A
c	22	228.4	2.3	601	4	US-09-949-016-63770	Sequence 63770, A
	23	227.6	2.3	92334	4	US-09-949-016-13920	Sequence 13920, A
	24	227.6	2.3	92363	4	US-09-949-016-12146	Sequence 12146, A
c	25	227.2	2.3	601	4	US-09-949-016-26018	Sequence 26018, A
c	26	227.2	2.3	601	4	US-09-949-016-74620	Sequence 74620, A
	27	227.2	2.3	88906	4	US-09-949-016-17468	Sequence 17468, A
c	28	226.2	2.3	27727	4	US-09-949-016-15737	Sequence 15737, A
	29	226.2	2.3	101951	4	US-09-949-016-15648	Sequence 15648, A
	30	225.2	2.3	53737	4	US-09-949-016-16197	Sequence 16197, A
	31	225.2	2.3	462589	4	US-09-949-016-12900	Sequence 12900, A
	32	225.2	2.3	476044	4	US-09-949-016-12412	Sequence 12412, A
	33	224.8	2.2	222691	4	US-09-949-016-11762	Sequence 11762, A
	34	224.8	2.2	222697	4	US-09-949-016-15842	Sequence 15842, A
	35	224.2	2.2	298	4	US-09-513-999C-11828	Sequence 11828, A
	36	224.2	2.2	66247	4	US-09-949-016-16009	Sequence 16009, A
c	37	224.2	2.2	390416	4	US-09-949-016-16923	Sequence 16923, A
	38	224	2.2	66804	4	US-09-740-041-3	Sequence 3, Appli
c	39	223.6	2.2	161124	4	US-09-949-016-11760	Sequence 11760, A
	40	223.6	2.2	176006	4	US-09-949-016-16804	Sequence 16804, A
	41	223.6	2.2	253375	4	US-09-949-016-12849	Sequence 12849, A
	42	223.2	2.2	601	4	US-09-949-016-118216	Sequence 118216,
	43	222.8	2.2	601	4	US-09-949-016-130896	Sequence 130896,
	44	222.8	2.2	36203	4	US-09-949-016-15430	Sequence 15430, A
c	45	222.8	2.2	90472	4	US-09-949-016-14038	Sequence 14038, A

ALIGNMENTS

RESULT 1
US-08-892-695-9
; Sequence 9, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892,695A
; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 10365
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Genomic
; OTHER INFORMATION: Sequence encoding ZABC1
; Patent No. 6808878
US-08-892-695-9

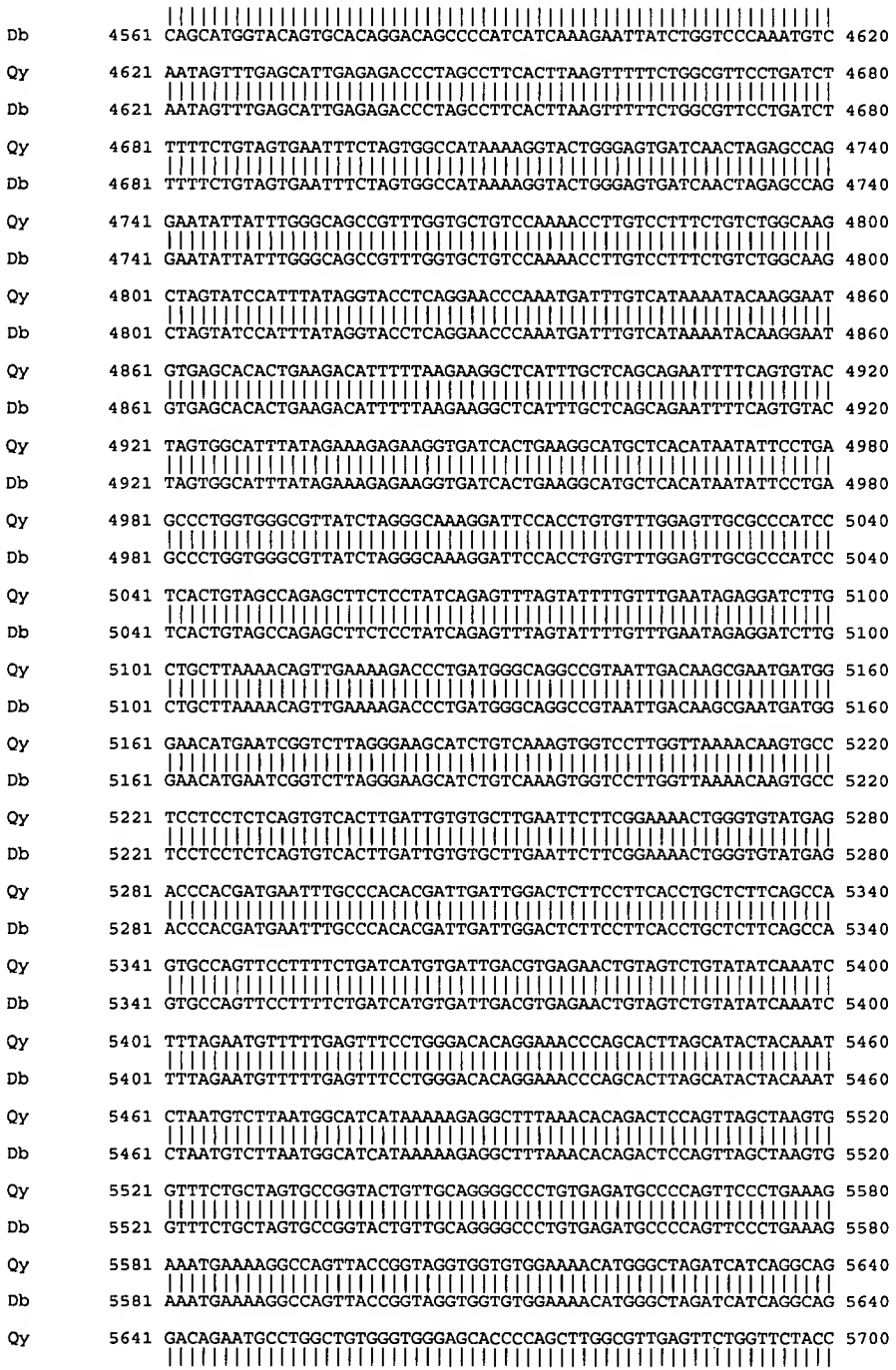
Query Match 100.0%; Score 10000; DB 4; Length 10365;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 10000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

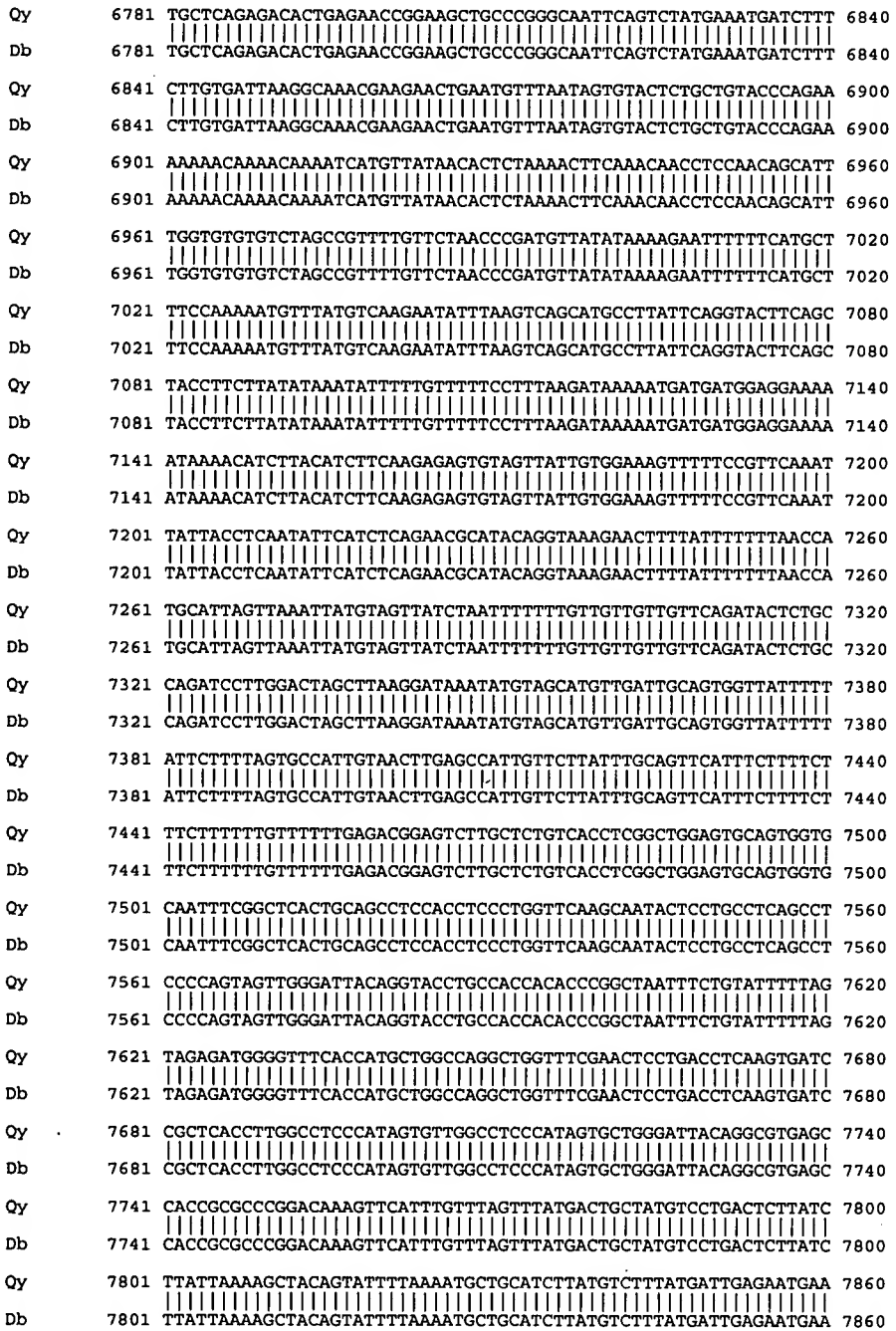
```
Qy          1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 60

Qy         61 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 120

Qy        121 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 180

Qy        181 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 240
```

```
Db      181 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 240
Qy      241 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300
Db      241 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300
Qy      301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 360
Db      301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 360
Qy      361 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 420
Db      361 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 420
Qy      421 AATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTTCAAAT 480
Db      421 AATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTTCAAAT 480
Qy      481 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAAAACTCCAC 540
Db      481 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAAAACTCCAC 540
Qy      541 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 600
Db      541 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 600
Qy      601 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 660
Db      601 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 660
Qy      661 TTTGATAACTGTTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 720
Db      661 TTTGATAACTGTTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 720
Qy      721 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 780
Db      721 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 780
Qy      781 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 840
Db      781 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 840
Qy      841 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 900
Db      841 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 900
Qy      901 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 960
Db      901 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 960
Qy      961 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 1020
Db      961 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 1020
Qy     1021 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 1080
Db     1021 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 1080
Qy     1081 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 1140
Db     1081 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 1140
Qy     1141 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 1200
Db     1141 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 1200
Qy     1201 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 1260
Db     1201 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 1260
Qy     1261 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 1320
```

```
Db     1261 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 1320
Qy     1321 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 1380
Db     1321 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 1380
Qy     1381 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 1440
Db     1381 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 1440
Qy     1441 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 1500
Db     1441 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 1500
Qy     1501 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATGA 1560
Db     1501 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATGA 1560
Qy     1561 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 1620
Db     1561 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 1620
Qy     1621 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 1680
Db     1621 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 1680
Qy     1681 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 1740
Db     1681 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 1740
Qy     1741 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 1800
Db     1741 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 1800
Qy     1801 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
Db     1801 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
Qy     1861 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 1920
Db     1861 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 1920
Qy     1921 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 1980
Db     1921 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 1980
Qy     1981 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 2040
Db     1981 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 2040
Qy     2041 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 2100
Db     2041 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 2100
Qy     2101 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 2160
Db     2101 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 2160
Qy     2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 2220
Db     2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 2220
Qy     2221 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 2280
Db     2221 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 2280
Qy     2281 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 2340
Db     2281 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 2340
Qy     2341 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTCT 2400
Db     2341 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTCT 2400
```

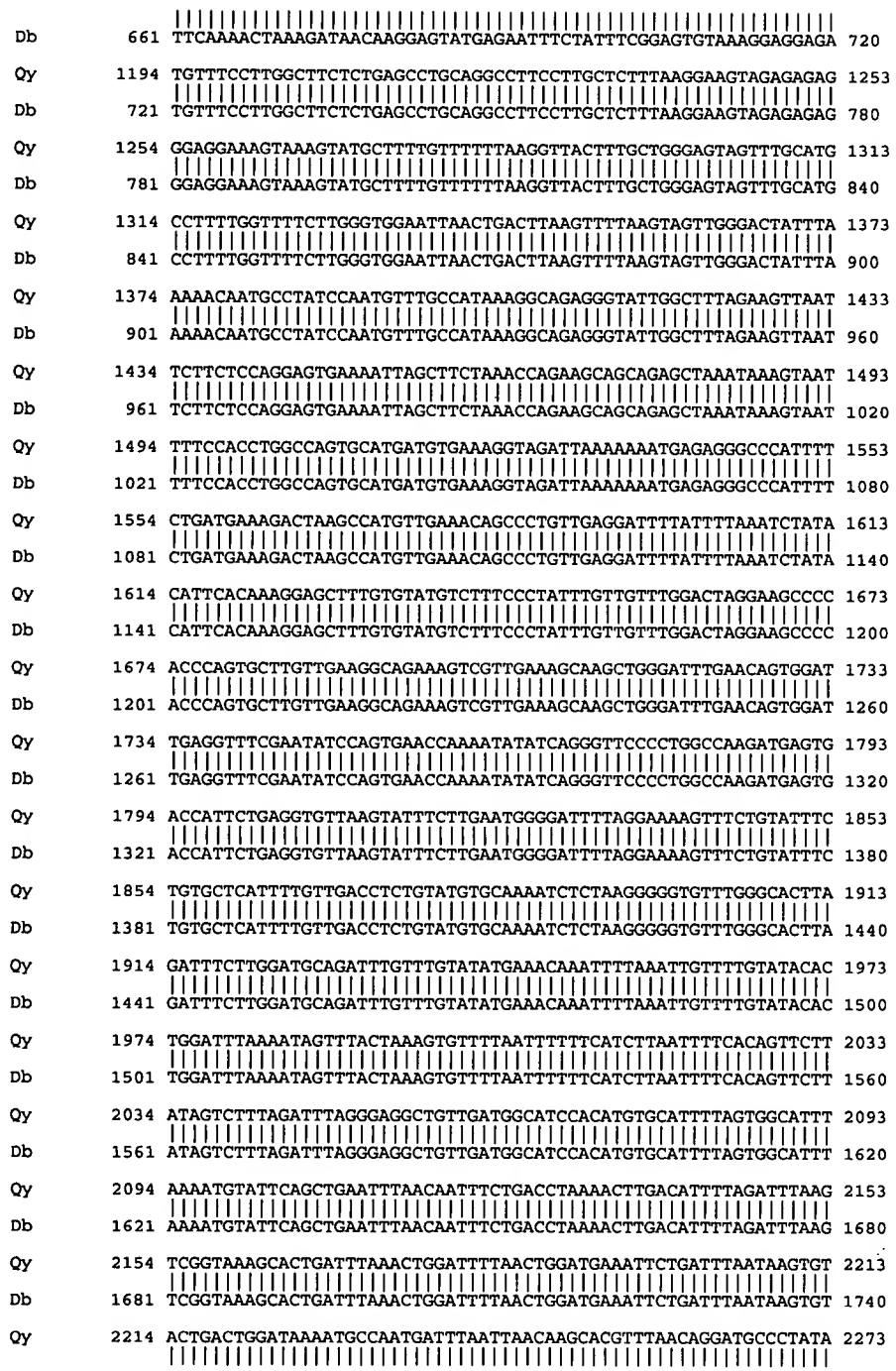
```
Qy      2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 2460

Qy      2461 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2461 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520

Qy      2521 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2521 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2580

Qy      2581 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2581 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 2640

Qy      2641 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2641 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2700

Qy      2701 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2701 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2760

Qy      2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2820

Qy      2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2880

Qy      2881 TTCCATTCCGAGCTACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2881 TTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 2940

Qy      2941 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2941 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 3000

Qy      3001 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3001 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 3060

Qy      3061 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3061 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 3120

Qy      3121 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3121 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 3180

Qy      3181 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3181 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 3240

Qy      3241 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3241 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 3300

Qy      3301 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3301 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 3360

Qy      3361 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3361 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 3420

Qy      3421 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3421 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 3480

Qy      3481 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3481 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 3540

Qy      3541 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3541 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 3600

Qy      3601 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3601 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 3660

Qy      3661 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3661 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 3720

Qy      3721 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3721 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 3780

Qy      3781 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 3840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3781 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 3840

Qy      3841 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 3900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3841 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 3900

Qy      3901 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 3960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3901 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 3960

Qy      3961 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 4020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3961 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 4020

Qy      4021 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 4080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4021 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 4080

Qy      4081 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 4140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4081 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 4140

Qy      4141 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 4200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4141 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 4200

Qy      4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 4260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 4260

Qy      4261 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 4320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4261 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 4320

Qy      4321 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 4380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4321 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 4380

Qy      4381 TCACAAAACACACTTTAAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 4440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4381 TCACAAAACACACTTTAAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 4440

Qy      4441 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 4500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4441 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 4500

Qy      4501 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 4560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4501 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 4560

Qy      4561 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 4620
```

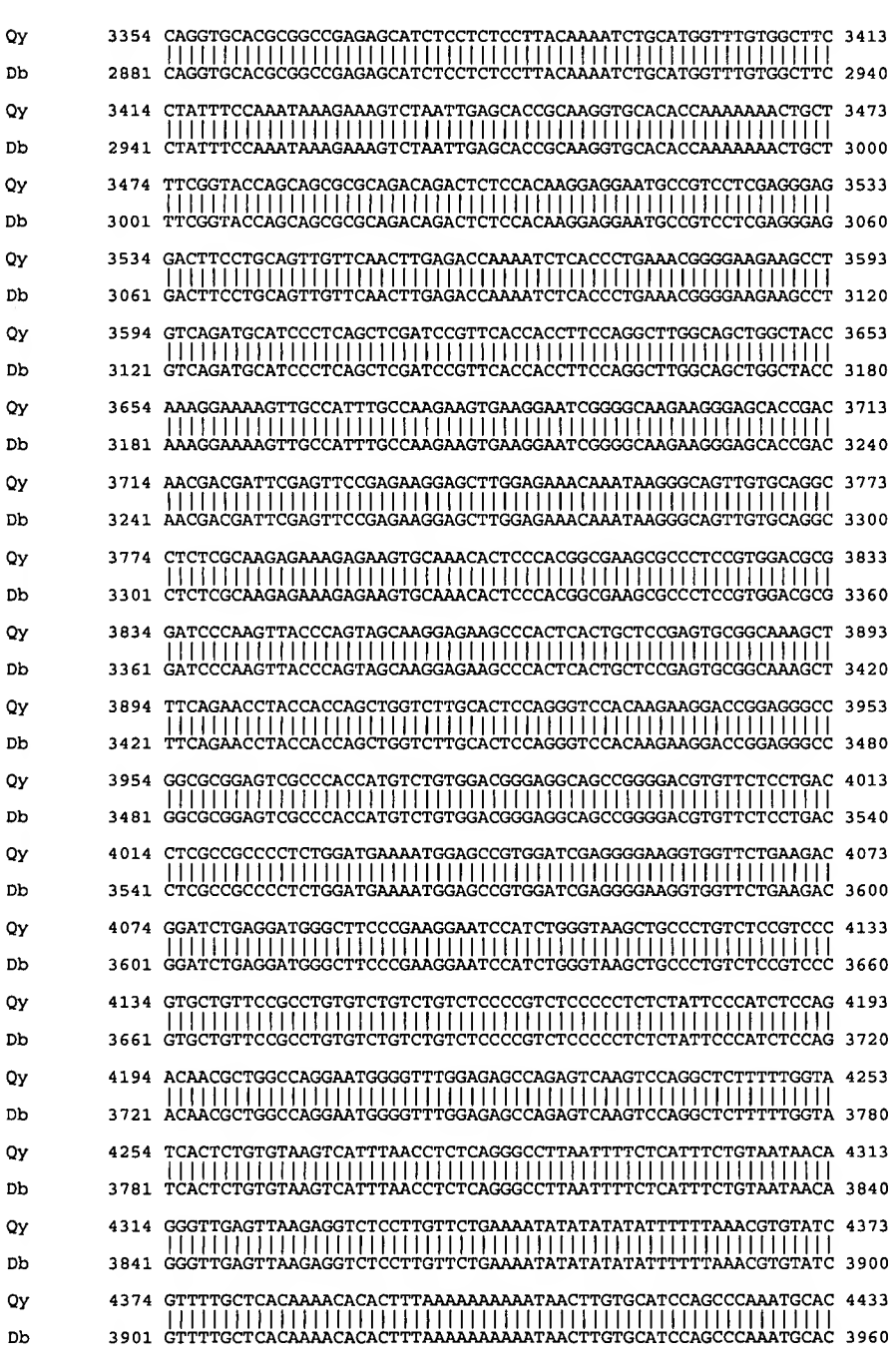

Db 4561 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 4620
Qy 4621 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 4680
Db 4621 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 4680
Qy 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATCAACTAGAGCCAG 4740
Db 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATCAACTAGAGCCAG 4740
Qy 4741 GAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGCAAG 4800
Db 4741 GAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGCAAG 4800
Qy 4801 CTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGGAAT 4860
Db 4801 CTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGGAAT 4860
Qy 4861 GTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTGTAC 4920
Db 4861 GTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTGTAC 4920
Qy 4921 TAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCCTGA 4980
Db 4921 TAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCCTGA 4980
Qy 4981 GCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCATCC 5040
Db 4981 GCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCATCC 5040
Qy 5041 TCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATCTTG 5100
Db 5041 TCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATCTTG 5100
Qy 5101 CTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGATGG 5160
Db 5101 CTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGATGG 5160
Qy 5161 GAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGTGCC 5220
Db 5161 GAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGTGCC 5220
Qy 5221 TCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTATGAG 5280
Db 5221 TCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTATGAG 5280
Qy 5281 ACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAGCCA 5340
Db 5281 ACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAGCCA 5340
Qy 5341 GTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAAATC 5400
Db 5341 GTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAAATC 5400
Qy 5401 TTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAAT 5460
Db 5401 TTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAAT 5460
Qy 5461 CTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAAGTG 5520
Db 5461 CTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAAGTG 5520
Qy 5521 GTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGAAAG 5580
Db 5521 GTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGAAAG 5580
Qy 5581 AAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGGCAG 5640
Db 5581 AAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGGCAG 5640
Qy 5641 GACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCTACC 5700
Db 5641 GACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCTACC 5700
Qy 5701 ACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCCCTG 5760
Db 5701 ACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCCCTG 5760
Qy 5761 TTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGCCTA 5820
Db 5761 TTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGCCTA 5820
Qy 5821 TTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCTTGA 5880
Db 5821 TTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCTTGA 5880
Qy 5881 AAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGCATT 5940
Db 5881 AAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGCATT 5940
Qy 5941 TCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGTCGG 6000
Db 5941 TCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGTCGG 6000
Qy 6001 AAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAAAAA 6060
Db 6001 AAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAAAAA 6060
Qy 6061 AGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAGTCT 6120
Db 6061 AGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAGTCT 6120
Qy 6121 AGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTAAAT 6180
Db 6121 AGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTAAAT 6180
Qy 6181 TTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTGGAT 6240
Db 6181 TTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTGGAT 6240
Qy 6241 AGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTTTGG 6300
Db 6241 AGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTTTGG 6300
Qy 6301 GGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTTGAA 6360
Db 6301 GGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTTGAA 6360
Qy 6361 GTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGTCCT 6420
Db 6361 GTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGTCCT 6420
Qy 6421 ATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTGTCAG 6480
Db 6421 ATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTGTCAG 6480
Qy 6481 ATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTGCTA 6540
Db 6481 ATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTGCTA 6540
Qy 6541 AATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAAGAT 6600
Db 6541 AATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAAGAT 6600
Qy 6601 TAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCTTAT 6660
Db 6601 TAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCTTAT 6660
Qy 6661 TTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTATTC 6720
Db 6661 TTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTATTC 6720
Qy 6721 TGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCTGCC 6780
Db 6721 TGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCTGCC 6780

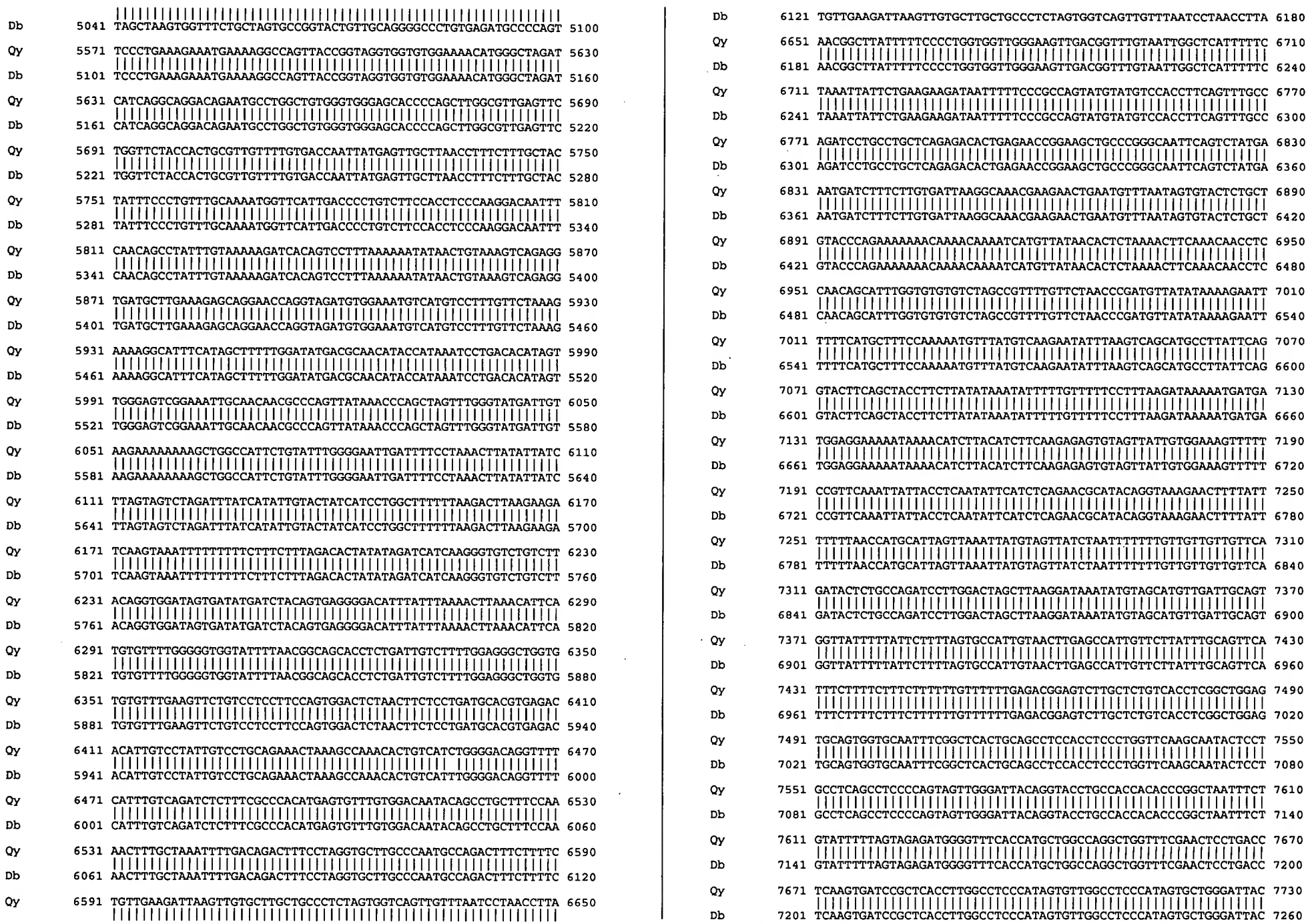

Qy 6781 TGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATCTTT 6840
Db 6781 TGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATCTTT 6840
Qy 6841 CTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCAGAA 6900
Db 6841 CTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCAGAA 6900
Qy 6901 AAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGCATT 6960
Db 6901 AAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGCATT 6960
Qy 6961 TGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCATGCT 7020
Db 6961 TGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCATGCT 7020
Qy 7021 TTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTCAGC 7080
Db 7021 TTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTCAGC 7080
Qy 7081 TACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGAAAA 7140
Db 7081 TACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGAAAA 7140
Qy 7141 ATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCAAAT 7200
Db 7141 ATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCAAAT 7200
Qy 7201 TATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAACCA 7260
Db 7201 TATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAACCA 7260
Qy 7261 TGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTCTGC 7320
Db 7261 TGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTCTGC 7320
Qy 7321 CAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATTTTT 7380
Db 7321 CAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATTTTT 7380
Qy 7381 ATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTTTCT 7440
Db 7381 ATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTTTCT 7440
Qy 7441 TTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTGGTG 7500
Db 7441 TTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTGGTG 7500
Qy 7501 CAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAGCCT 7560
Db 7501 CAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAGCCT 7560
Qy 7561 CCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTTTAG 7620
Db 7561 CCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTTTAG 7620
Qy 7621 TAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATC 7680
Db 7621 TAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATC 7680
Qy 7681 CGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTGAGC 7740
Db 7681 CGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTGAGC 7740
Qy 7741 CACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTTATC 7800
Db 7741 CACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTTATC 7800
Qy 7801 TTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAATGAA 7860
Db 7801 TTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAATGAA 7860
Qy 7861 ATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTAGGA 7920
Db 7861 ATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTAGGA 7920
Qy 7921 GGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCAAGT 7980
Db 7921 GGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCAAGT 7980
Qy 7981 TATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATATTTC 8040
Db 7981 TATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATATTTC 8040
Qy 8041 CCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATTGAA 8100
Db 8041 CCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATTGAA 8100
Qy 8101 GATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGAGAA 8160
Db 8101 GATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGAGAA 8160
Qy 8161 ATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAATAA 8220
Db 8161 ATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAATAA 8220
Qy 8221 ATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTTCTT 8280
Db 8221 ATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTTCTT 8280
Qy 8281 GCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACA 8340
Db 8281 GCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACA 8340
Qy 8341 TCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAA 8400
Db 8341 TCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAA 8400
Qy 8401 GTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCG 8460
Db 8401 GTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCG 8460
Qy 8461 CAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCT 8520
Db 8461 CAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCT 8520
Qy 8521 GCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTC 8580
Db 8521 GCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTC 8580
Qy 8581 TCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGAT 8640
Db 8581 TCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGAT 8640
Qy 8641 AAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTT 8700
Db 8641 AAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTT 8700
Qy 8701 GAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGAT 8760
Db 8701 GAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGAT 8760
Qy 8761 GGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCT 8820
Db 8761 GGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCT 8820
Qy 8821 GACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGG 8880
Db 8821 GACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGG 8880
Qy 8881 GCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGT 8940
Db 8881 GCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGT 8940
Qy 8941 CCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAG 9000

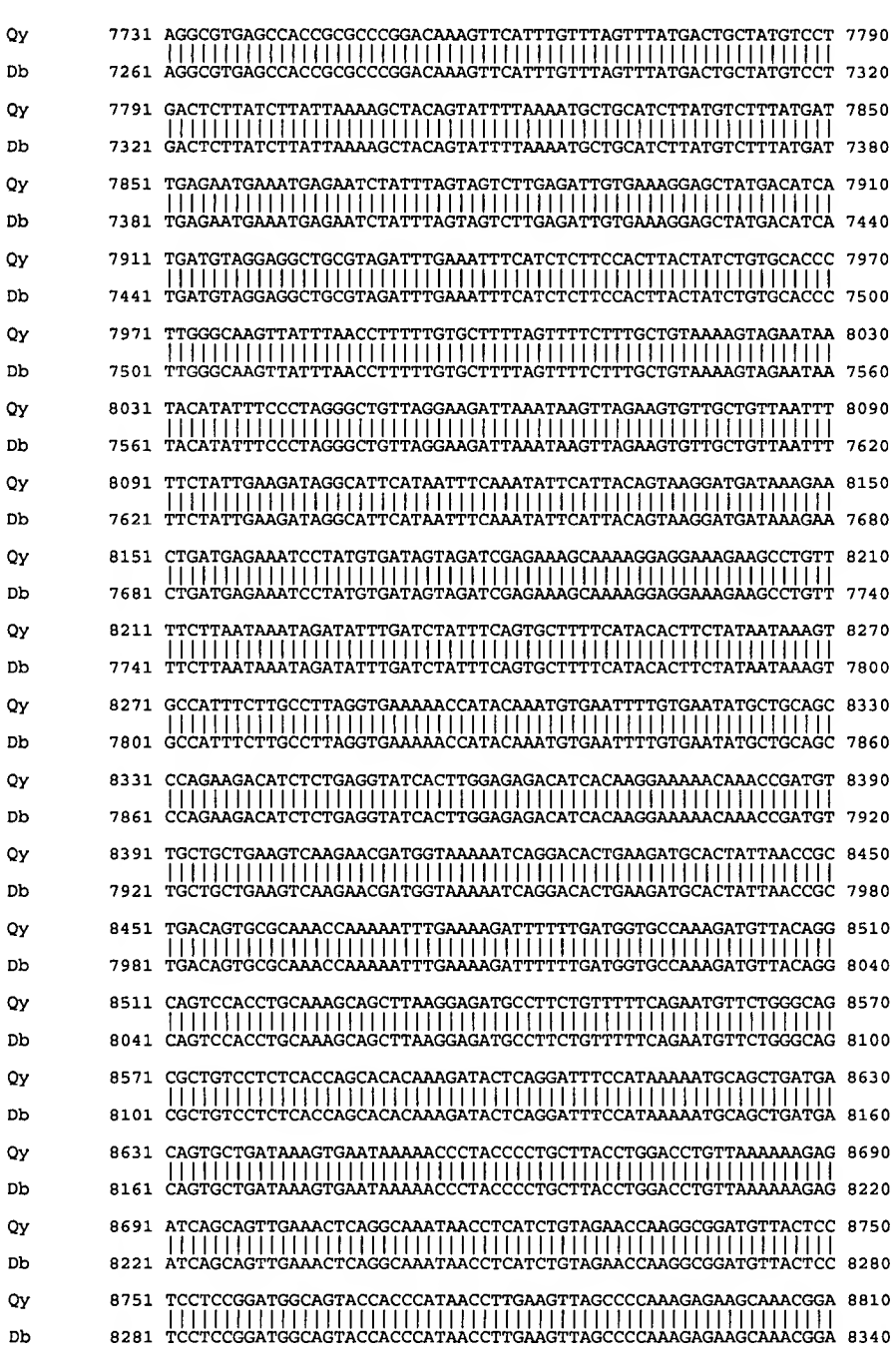
```
Db       8941 CCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAG 9000

Qy       9001 CATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGA 9060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9001 CATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGA 9060

Qy       9061 CGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGT 9120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9061 CGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGT 9120

Qy       9121 AAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGG 9180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9121 AAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGG 9180

Qy       9181 AAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACT 9240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9181 AAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACT 9240

Qy       9241 TTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCC 9300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9241 TTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCC 9300

Qy       9301 GCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGAT 9360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9301 GCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGAT 9360

Qy       9361 AAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACC 9420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9361 AAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACC 9420

Qy       9421 CTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGG 9480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9421 CTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGG 9480

Qy       9481 GCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTT 9540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9481 GCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTT 9540

Qy       9541 GGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCC 9600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9541 GGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCC 9600

Qy       9601 GGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACA 9660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9601 GGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACA 9660

Qy       9661 TCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTC 9720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9661 TCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTC 9720

Qy       9721 CTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGC 9780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9721 CTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGC 9780

Qy       9781 TCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAA 9840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9781 TCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAA 9840

Qy       9841 GGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAATCCA 9900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9841 GGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAATCCA 9900

Qy       9901 GGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGTTCT 9960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9901 GGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGTTCT 9960

Qy       9961 TGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000
             |||||||||||||||||||||||||||||||||||||||
Db       9961 TGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000


RESULT 2
```

```
US-09-949-016-12604
; Sequence 12604, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12604
;   LENGTH: 20022
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12604

  Query Match             95.1%;  Score 9509.2;  DB 4;  Length 20022;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 9524;  Conservative    0;  Mismatches    3;  Indels    3;  Gaps    1;

Qy        474 TTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAACCAAAAA 533
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAACCAAAAA 60

Qy        534 ACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATC 593
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATC 120

Qy        594 TAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTT 653
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTT 180

Qy        654 GTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCT 713
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCT 240

Qy        714 GGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGT 773
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGT 300

Qy        774 ATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATG 833
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATG 360

Qy        834 TGTAGCTCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAA 893
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTAGCTCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAA 420

Qy        894 CCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTT 953
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTT 480

Qy        954 GTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCC 1013
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCC 540

Qy       1014 CATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTT 1073
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTT 600

Qy       1074 TACCCTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTA 1133
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TACCCTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTA 660

Qy       1134 TTCAAAACTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGA 1193
```

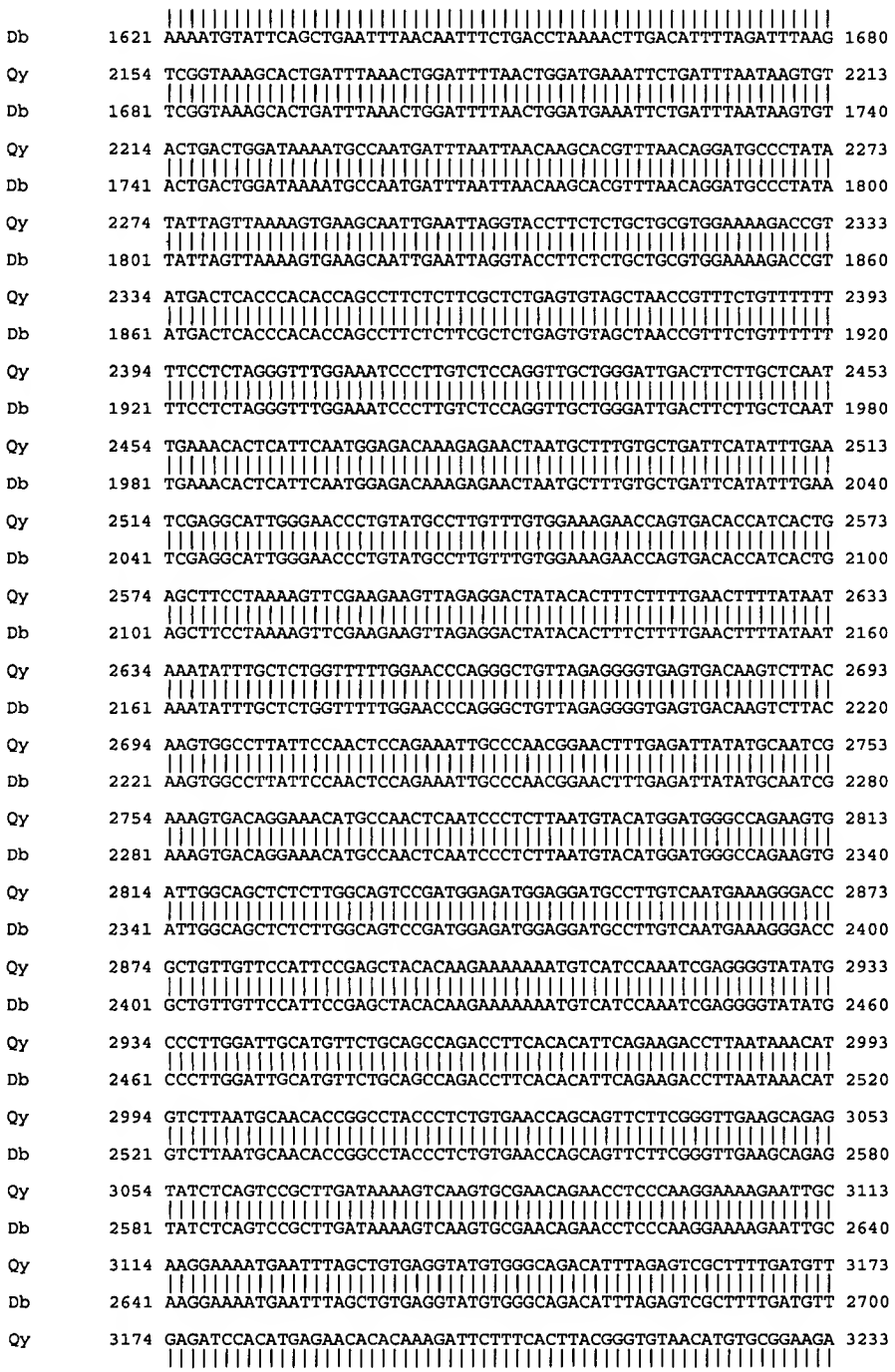

Db 661 TTCAAAACTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGA 720
Qy 1194 TGTTTCCTTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAG 1253
Db 721 TGTTTCCTTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAG 780
Qy 1254 GGAGGAAAGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATG 1313
Db 781 GGAGGAAAGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATG 840
Qy 1314 CCTTTTGGTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTA 1373
Db 841 CCTTTTGGTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTA 900
Qy 1374 AAAACAATGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAAT 1433
Db 901 AAAACAATGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAAT 960
Qy 1434 TCTTCTCCAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAAT 1493
Db 961 TCTTCTCCAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAAT 1020
Qy 1494 TTTCCACCTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTT 1553
Db 1021 TTTCCACCTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTT 1080
Qy 1554 CTGATGAAAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATA 1613
Db 1081 CTGATGAAAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATA 1140
Qy 1614 CATTCACAAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCC 1673
Db 1141 CATTCACAAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCC 1200
Qy 1674 ACCCAGTGCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGAT 1733
Db 1201 ACCCAGTGCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGAT 1260
Qy 1734 TGAGGTTTCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTG 1793
Db 1261 TGAGGTTTCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTG 1320
Qy 1794 ACCATTCTGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTC 1853
Db 1321 ACCATTCTGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTC 1380
Qy 1854 TGTGCTCATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTA 1913
Db 1381 TGTGCTCATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTA 1440
Qy 1914 GATTTCTTGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACAC 1973
Db 1441 GATTTCTTGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACAC 1500
Qy 1974 TGGATTTAAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTT 2033
Db 1501 TGGATTTAAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTT 1560
Qy 2034 ATAGTCTTTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTT 2093
Db 1561 ATAGTCTTTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTT 1620
Qy 2094 AAAATGTATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAG 2153
Db 1621 AAAATGTATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAG 1680
Qy 2154 TCGGTAAAGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGT 2213
Db 1681 TCGGTAAAGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGT 1740
Qy 2214 ACTGACTGGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATA 2273
Db 1741 ACTGACTGGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATA 1800
Qy 2274 TATTAGTTAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGT 2333
Db 1801 TATTAGTTAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGT 1860
Qy 2334 ATGACTCACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTT 2393
Db 1861 ATGACTCACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTT 1920
Qy 2394 TTCCTCTAGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAAT 2453
Db 1921 TTCCTCTAGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAAT 1980
Qy 2454 TGAAACACTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAA 2513
Db 1981 TGAAACACTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAA 2040
Qy 2514 TCGAGGCATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTG 2573
Db 2041 TCGAGGCATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTG 2100
Qy 2574 AGCTTCCTAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAAT 2633
Db 2101 AGCTTCCTAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAAT 2160
Qy 2634 AAATATTTGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTAC 2693
Db 2161 AAATATTTGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTAC 2220
Qy 2694 AAGTGGCCTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCG 2753
Db 2221 AAGTGGCCTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCG 2280
Qy 2754 AAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTG 2813
Db 2281 AAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTG 2340
Qy 2814 ATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACC 2873
Db 2341 ATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACC 2400
Qy 2874 GCTGTTGTTCCATTCCGAGCTACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATG 2933
Db 2401 GCTGTTGTTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATG 2460
Qy 2934 CCCTTGGATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACAT 2993
Db 2461 CCCTTGGATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACAT 2520
Qy 2994 GTCTTAATGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAG 3053
Db 2521 GTCTTAATGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAG 2580
Qy 3054 TATCTCAGTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGC 3113
Db 2581 TATCTCAGTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGC 2640
Qy 3114 AAGGAAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTT 3173
Db 2641 AAGGAAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTT 2700
Qy 3174 GAGATCCACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGA 3233
Db 2701 GAGATCCACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGA 2760
Qy 3234 AGATTCAAGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGG 3293
Db 2761 AGATTCAAGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGG 2820
Qy 3294 GCCAGAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTC 3353
Db 2821 GCCAGAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTC 2880


Qy 3354 CAGGTGCACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTC 3413
Db 2881 CAGGTGCACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTC 2940
Qy 3414 CTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCT 3473
Db 2941 CTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCT 3000
Qy 3474 TTCGGTACCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAG 3533
Db 3001 TTCGGTACCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAG 3060
Qy 3534 GACTTCCTGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCT 3593
Db 3061 GACTTCCTGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCT 3120
Qy 3594 GTCAGATGCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACC 3653
Db 3121 GTCAGATGCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACC 3180
Qy 3654 AAAGGAAAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGAC 3713
Db 3181 AAAGGAAAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGAC 3240
Qy 3714 AACGACGATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGC 3773
Db 3241 AACGACGATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGC 3300
Qy 3774 CTCTCGCAAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCG 3833
Db 3301 CTCTCGCAAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCG 3360
Qy 3834 GATCCCAAGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCT 3893
Db 3361 GATCCCAAGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCT 3420
Qy 3894 TTCAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCC 3953
Db 3421 TTCAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCC 3480
Qy 3954 GGCGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGAC 4013
Db 3481 GGCGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGAC 3540
Qy 4014 CTCGCCGCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGAC 4073
Db 3541 CTCGCCGCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGAC 3600
Qy 4074 GGATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCC 4133
Db 3601 GGATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCC 3660
Qy 4134 GTGCTGTTCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAG 4193
Db 3661 GTGCTGTTCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAG 3720
Qy 4194 ACAACGCTGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTA 4253
Db 3721 ACAACGCTGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTA 3780
Qy 4254 TCACTCTGTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACA 4313
Db 3781 TCACTCTGTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACA 3840
Qy 4314 GGGTTGAGTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATC 4373
Db 3841 GGGTTGAGTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATC 3900
Qy 4374 GTTTTGCTCACAAAACACACTTTAAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCAC 4433
Db 3901 GTTTTGCTCACAAAACACACTTTAAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCAC 3960
Qy 4434 TGCTTCTTAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTT 4493
Db 3961 TGCTTCTTAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTT 4020
Qy 4494 TGGTTGTCATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGA 4553
Db 4021 TGGTTGTCATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGA 4080
Qy 4554 CACTGCTCAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCC 4613
Db 4081 CACTGCTCAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCC 4140
Qy 4614 AAATGTCAATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTC 4673
Db 4141 AAATGTCAATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTC 4200
Qy 4674 CTGATCTTTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAG---TGATCAA 4730
Db 4201 CTGATCTTTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATGATCAA 4260
Qy 4731 CTAGAGCCAGGAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCT 4790
Db 4261 CTAGAGCCAGGAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCT 4320
Qy 4791 GTCTGGCAAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAA 4850
Db 4321 GTCTGGCAAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAA 4380
Qy 4851 TACAAGGAATGTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATT 4910
Db 4381 TACAAGGAATGTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATT 4440
Qy 4911 TTCAGTGTACTAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATA 4970
Db 4441 TTCAGTGTACTAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATA 4500
Qy 4971 ATATTCCTGAGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTT 5030
Db 4501 ATATTCCTGAGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTT 4560
Qy 5031 GCGCCCATCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATA 5090
Db 4561 GCGCCCATCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATA 4620
Qy 5091 GAGGATCTTGCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAG 5150
Db 4621 GAGGATCTTGCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAG 4680
Qy 5151 CGAATGATGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAA 5210
Db 4681 CGAATGATGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTCGGTTAA 4740
Qy 5211 AACAAGTGCCTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTG 5270
Db 4741 AACAAGTGCCTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTG 4800
Qy 5271 GGTGTATGAGACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGC 5330
Db 4801 GGTGTATGAGACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGC 4860
Qy 5331 TCTTCAGCCAGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGT 5390
Db 4861 TCTTCAGCCAGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGT 4920
Qy 5391 ATATCAAATCTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCA 5450
Db 4921 ATATCAAATCTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCA 4980
Qy 5451 TACTACAAATCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGT 5510
Db 4981 TACTACAAATCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGT 5040
Qy 5511 TAGCTAAGTGGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGT 5570

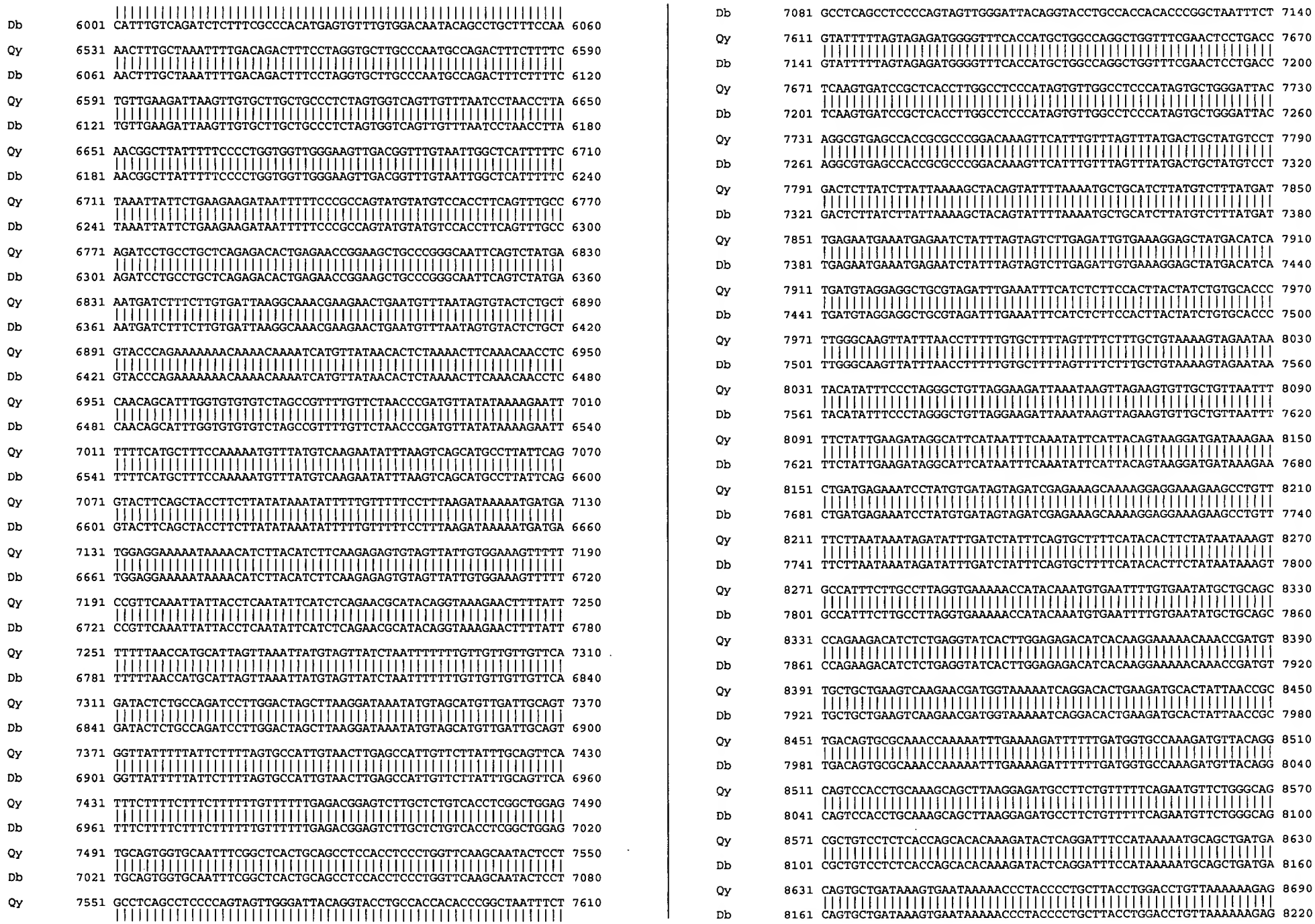

Db 5041 TAGCTAAGTGGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGT 5100
Qy 5571 TCCCTGAAAGAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGAT 5630
Db 5101 TCCCTGAAAGAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGAT 5160
Qy 5631 CATCAGGCAGGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTC 5690
Db 5161 CATCAGGCAGGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTC 5220
Qy 5691 TGGTTCTACCACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTAC 5750
Db 5221 TGGTTCTACCACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTAC 5280
Qy 5751 TATTTCCCTGTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTT 5810
Db 5281 TATTTCCCTGTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTT 5340
Qy 5811 CAACAGCCTATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGG 5870
Db 5341 CAACAGCCTATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGG 5400
Qy 5871 TGATGCTTGAAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAG 5930
Db 5401 TGATGCTTGAAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAG 5460
Qy 5931 AAAAGGCATTTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGT 5990
Db 5461 AAAAGGCATTTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGT 5520
Qy 5991 TGGGAGTCGGAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGT 6050
Db 5521 TGGGAGTCGGAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGT 5580
Qy 6051 AAGAAAAAAAAAGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATC 6110
Db 5581 AAGAAAAAAAAAGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATC 5640
Qy 6111 TTAGTAGTCTAGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGA 6170
Db 5641 TTAGTAGTCTAGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGA 5700
Qy 6171 TCAAGTAAATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTT 6230
Db 5701 TCAAGTAAATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTT 5760
Qy 6231 ACAGGTGGATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCA 6290
Db 5761 ACAGGTGGATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCA 5820
Qy 6291 TGTGTTTTGGGGGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTG 6350
Db 5821 TGTGTTTTGGGGGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTG 5880
Qy 6351 TGTGTTTGAAGTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGAC 6410
Db 5881 TGTGTTTGAAGTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGAC 5940
Qy 6411 ACATTGTCCTATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTT 6470
Db 5941 ACATTGTCCTATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATTTGGGGACAGGTTTT 6000
Qy 6471 CATTTGTCAGATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAA 6530
Db 6001 CATTTGTCAGATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAA 6060
Qy 6531 AACTTTGCTAAATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTC 6590
Db 6061 AACTTTGCTAAATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTC 6120
Qy 6591 TGTTGAAGATTAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTA 6650
Db 6121 TGTTGAAGATTAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTA 6180
Qy 6651 AACGGCTTATTTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTC 6710
Db 6181 AACGGCTTATTTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTC 6240
Qy 6711 TAAATTATTCTGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCC 6770
Db 6241 TAAATTATTCTGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCC 6300
Qy 6771 AGATCCTGCCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGA 6830
Db 6301 AGATCCTGCCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGA 6360
Qy 6831 AATGATCTTTCTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCT 6890
Db 6361 AATGATCTTTCTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCT 6420
Qy 6891 GTACCCAGAAAAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTC 6950
Db 6421 GTACCCAGAAAAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTC 6480
Qy 6951 CAACAGCATTTGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATT 7010
Db 6481 CAACAGCATTTGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATT 6540
Qy 7011 TTTTCATGCTTTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAG 7070
Db 6541 TTTTCATGCTTTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAG 6600
Qy 7071 GTACTTCAGCTACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGA 7130
Db 6601 GTACTTCAGCTACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGA 6660
Qy 7131 TGGAGGAAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTT 7190
Db 6661 TGGAGGAAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTT 6720
Qy 7191 CCGTTCAAATTATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATT 7250
Db 6721 CCGTTCAAATTATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATT 6780
Qy 7251 TTTTTAACCATGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCA 7310
Db 6781 TTTTTAACCATGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCA 6840
Qy 7311 GATACTCTGCCAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGT 7370
Db 6841 GATACTCTGCCAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGT 6900
Qy 7371 GGTTATTTTTATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCA 7430
Db 6901 GGTTATTTTTATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCA 6960
Qy 7431 TTTCTTTTCTTTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAG 7490
Db 6961 TTTCTTTTCTTTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAG 7020
Qy 7491 TGCAGTGGTGCAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCT 7550
Db 7021 TGCAGTGGTGCAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCT 7080
Qy 7551 GCCTCAGCCTCCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCT 7610
Db 7081 GCCTCAGCCTCCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCT 7140
Qy 7611 GTATTTTTAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACC 7670
Db 7141 GTATTTTTAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACC 7200
Qy 7671 TCAAGTGATCCGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTAC 7730
Db 7201 TCAAGTGATCCGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTAC 7260


Qy 7731 AGGCGTGAGCCACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCT 7790
Db 7261 AGGCGTGAGCCACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCT 7320
Qy 7791 GACTCTTATCTTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGAT 7850
Db 7321 GACTCTTATCTTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGAT 7380
Qy 7851 TGAGAATGAAATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCA 7910
Db 7381 TGAGAATGAAATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCA 7440
Qy 7911 TGATGTAGGAGGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCC 7970
Db 7441 TGATGTAGGAGGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCC 7500
Qy 7971 TTGGGCAAGTTATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAA 8030
Db 7501 TTGGGCAAGTTATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAA 7560
Qy 8031 TACATATTTCCCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTT 8090
Db 7561 TACATATTTCCCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTT 7620
Qy 8091 TTCTATTGAAGATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAA 8150
Db 7621 TTCTATTGAAGATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAA 7680
Qy 8151 CTGATGAGAAATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTT 8210
Db 7681 CTGATGAGAAATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTT 7740
Qy 8211 TTCTTAATAAATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGT 8270
Db 7741 TTCTTAATAAATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGT 7800
Qy 8271 GCCATTTCTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGC 8330
Db 7801 GCCATTTCTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGC 7860
Qy 8331 CCAGAAGACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGT 8390
Db 7861 CCAGAAGACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGT 7920
Qy 8391 TGCTGCTGAAGTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGC 8450
Db 7921 TGCTGCTGAAGTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGC 7980
Qy 8451 TGACAGTGCGCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGG 8510
Db 7981 TGACAGTGCGCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGG 8040
Qy 8511 CAGTCCACCTGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAG 8570
Db 8041 CAGTCCACCTGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAG 8100
Qy 8571 CGCTGTCCTCTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGA 8630
Db 8101 CGCTGTCCTCTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGA 8160
Qy 8631 CAGTGCTGATAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAG 8690
Db 8161 CAGTGCTGATAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAG 8220
Qy 8691 ATCAGCAGTTGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCC 8750
Db 8221 ATCAGCAGTTGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCC 8280
Qy 8751 TCCTCCGGATGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGA 8810
Db 8281 TCCTCCGGATGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGA 8340
Qy 8811 GACCGCAGCTGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTT 8870
Db 8341 GACCGCAGCTGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTT 8400
Qy 8871 ATCCGTGGGGGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAG 8930
Db 8401 ATCCGTGGGGGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAG 8460
Qy 8931 TATCACCTGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCA 8990
Db 8461 TATCACCTGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCA 8520
Qy 8991 GAGACTGGAGCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCT 9050
Db 8521 GAGACTGGAGCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCT 8580
Qy 9051 TAGAAGTCGACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCC 9110
Db 8581 TAGAAGTCGACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCTC 8640
Qy 9111 TAGTTTCTGTAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATC 9170
Db 8641 TAGTTTCTGTAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATC 8700
Qy 9171 TGCGAAGGGGAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGA 9230
Db 8701 TGCGAAGGGGAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGA 8760
Qy 9231 CTCTAGCACTTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGT 9290
Db 8761 CTCTAGCACTTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGT 8820
Qy 9291 CCAAGGGGCCGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCC 9350
Db 8821 CCAAGGGGCCGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCC 8880
Qy 9351 TGCACCGGATAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTC 9410
Db 8881 TGCACCGGATAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTC 8940
Qy 9411 TCAGCCCACCCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGA 9470
Db 8941 TCAGCCCACCCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGA 9000
Qy 9471 CAGCCCGTGGGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGC 9530
Db 9001 CAGCCCGTGGGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGC 9060
Qy 9531 AGCAGAATTTGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGT 9590
Db 9061 AGCAGAATTTGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGT 9120
Qy 9591 TGACCAGCCCGGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAG 9650
Db 9121 TGACCAGCCCGGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAG 9180
Qy 9651 AGGCATCACATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAA 9710
Db 9181 AGGCATCACATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAA 9240
Qy 9711 ACCAAGGTTCCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCC 9770
Db 9241 ACCAAGGTTCCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCC 9300
Qy 9771 CTATGGTGGCTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTC 9830
Db 9301 CTATGGTGGCTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTC 9360
Qy 9831 GACGTTAGAAGGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAAT 9890
Db 9361 GACGTTAGAAGGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAAT 9420
Qy 9891 AGCTAATCCAGGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGC 9950

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        9421 AGCTAATCCAGGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGTAGGCAGC 9480

Qy        9951 CAGAAGTTCTTGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        9481 CAGAAGTTCTTGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 9530

RESULT 3
US-09-949-016-16004
; Sequence 16004, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16004
;   LENGTH: 20023
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16004

  Query Match             95.1%;  Score 9509.2;  DB 4;  Length 20023;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 9524;  Conservative    0;  Mismatches    3;  Indels    3;  Gaps    1;

Qy         474 TTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAACCAAAAA 533
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 TTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAACCAAAAA 60

Qy         534 ACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATC 593
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 ACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATC 120

Qy         594 TAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTT 653
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTT 180

Qy         654 GTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCT 713
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCT 240

Qy         714 GGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGT 773
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGT 300

Qy         774 ATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATG 833
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATG 360

Qy         834 TGTAGCTCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAA 893
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TGTAGCTCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAA 420

Qy         894 CCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTT 953
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTT 480

Qy         954 GTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCC 1013
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 GTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCC 540

Qy        1014 CATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTT 1073
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTT 600

Qy        1074 TACCCTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTA 1133
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 TACCCTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTA 660

Qy        1134 TTCAAAACTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGA 1193
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 TTCAAAACTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGA 720

Qy        1194 TGTTTCCTTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAG 1253
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 TGTTTCCTTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAG 780

Qy        1254 GGAGGAAAGTAAAGTATGCTTTTGTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATG 1313
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 GGAGGAAAGTAAAGTATGCTTTTGTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATG 840

Qy        1314 CCTTTTGGTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTA 1373
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 CCTTTTGGTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTA 900

Qy        1374 AAAACAATGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAAT 1433
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 AAAACAATGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAAT 960

Qy        1434 TCTTCTCCAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAAT 1493
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 TCTTCTCCAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAAT 1020

Qy        1494 TTTCCACCTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTT 1553
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 TTTCCACCTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTT 1080

Qy        1554 CTGATGAAAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATA 1613
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 CTGATGAAAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATA 1140

Qy        1614 CATTCACAAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCC 1673
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CATTCACAAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCC 1200

Qy        1674 ACCCAGTGCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGAT 1733
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 ACCCAGTGCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGAT 1260

Qy        1734 TGAGGTTTCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTG 1793
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 TGAGGTTTCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTG 1320

Qy        1794 ACCATTCTGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTC 1853
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 ACCATTCTGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTC 1380

Qy        1854 TGTGCTCATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTA 1913
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 TGTGCTCATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTA 1440

Qy        1914 GATTTCTTGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACAC 1973
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 GATTTCTTGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACAC 1500

Qy        1974 TGGATTTAAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTT 2033
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 TGGATTTAAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTT 1560

Qy        2034 ATAGTCTTTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTT 2093
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 ATAGTCTTTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTT 1620

Qy        2094 AAAATGTATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAG 2153
```


Db 1621 AAAATGTATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAG 1680
Qy 2154 TCGGTAAAGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGT 2213
Db 1681 TCGGTAAAGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGT 1740
Qy 2214 ACTGACTGGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATA 2273
Db 1741 ACTGACTGGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATA 1800
Qy 2274 TATTAGTTAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGT 2333
Db 1801 TATTAGTTAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGT 1860
Qy 2334 ATGACTCACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTT 2393
Db 1861 ATGACTCACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTT 1920
Qy 2394 TTCCTCTAGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAAT 2453
Db 1921 TTCCTCTAGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAAT 1980
Qy 2454 TGAAACACTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAA 2513
Db 1981 TGAAACACTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAA 2040
Qy 2514 TCGAGGCATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTG 2573
Db 2041 TCGAGGCATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTG 2100
Qy 2574 AGCTTCCTAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAAT 2633
Db 2101 AGCTTCCTAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAAT 2160
Qy 2634 AAATATTTGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTAC 2693
Db 2161 AAATATTTGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTAC 2220
Qy 2694 AAGTGGCCTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCG 2753
Db 2221 AAGTGGCCTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCG 2280
Qy 2754 AAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTG 2813
Db 2281 AAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTG 2340
Qy 2814 ATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACC 2873
Db 2341 ATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACC 2400
Qy 2874 GCTGTTGTTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATG 2933
Db 2401 GCTGTTGTTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATG 2460
Qy 2934 CCCTTGGATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACAT 2993
Db 2461 CCCTTGGATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACAT 2520
Qy 2994 GTCTTAATGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAG 3053
Db 2521 GTCTTAATGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAG 2580
Qy 3054 TATCTCAGTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGC 3113
Db 2581 TATCTCAGTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGC 2640
Qy 3114 AAGGAAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTT 3173
Db 2641 AAGGAAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTT 2700
Qy 3174 GAGATCCACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGA 3233
Db 2701 GAGATCCACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGA 2760
Qy 3234 AGATTCAAGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGG 3293
Db 2761 AGATTCAAGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGG 2820
Qy 3294 GCCAGAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTC 3353
Db 2821 GCCAGAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTC 2880
Qy 3354 CAGGTGCACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTC 3413
Db 2881 CAGGTGCACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTC 2940
Qy 3414 CTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCT 3473
Db 2941 CTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCT 3000
Qy 3474 TTCGGTACCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAG 3533
Db 3001 TTCGGTACCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAG 3060
Qy 3534 GACTTCCTGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCT 3593
Db 3061 GACTTCCTGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCT 3120
Qy 3594 GTCAGATGCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACC 3653
Db 3121 GTCAGATGCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACC 3180
Qy 3654 AAAGGAAAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGAC 3713
Db 3181 AAAGGAAAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGAC 3240
Qy 3714 AACGACGATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGC 3773
Db 3241 AACGACGATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGC 3300
Qy 3774 CTCTCGCAAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCG 3833
Db 3301 CTCTCGCAAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCG 3360
Qy 3834 GATCCCAAGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCT 3893
Db 3361 GATCCCAAGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCT 3420
Qy 3894 TTCAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCC 3953
Db 3421 TTCAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCC 3480
Qy 3954 GGCGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGAC 4013
Db 3481 GGCGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGAC 3540
Qy 4014 CTCGCCGCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGAC 4073
Db 3541 CTCGCCGCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGAC 3600
Qy 4074 GGATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCC 4133
Db 3601 GGATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCC 3660
Qy 4134 GTGCTGTTCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAG 4193
Db 3661 GTGCTGTTCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAG 3720
Qy 4194 ACAACGCTGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTA 4253
Db 3721 ACAACGCTGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTA 3780
Qy 4254 TCACTCTGTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACA 4313
Db 3781 TCACTCTGTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACA 3840

```
Qy       4314 GGGTTGAGTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATC 4373
Db       3841 GGGTTGAGTTAAGAGGTCTCCTTGTTCTGAAAATATATATATTTTTTAAACGTGTATC 3900

Qy       4374 GTTTTGCTCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCAC 4433
Db       3901 GTTTTGCTCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCAC 3960

Qy       4434 TGCTTCTTAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTT 4493
Db       3961 TGCTTCTTAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTT 4020

Qy       4494 TGGTTGTCATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGA 4553
Db       4021 TGGTTGTCATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGA 4080

Qy       4554 CACTGCTCAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCC 4613
Db       4081 CACTGCTCAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCC 4140

Qy       4614 AAATGTCAATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTC 4673
Db       4141 AAATGTCAATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTC 4200

Qy       4674 CTGATCTTTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAG---TGATCAA 4730
Db       4201 CTGATCTTTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATGATCAA 4260

Qy       4731 CTAGAGCCAGGAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCT 4790
Db       4261 CTAGAGCCAGGAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCT 4320

Qy       4791 GTCTGGCAAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAA 4850
Db       4321 GTCTGGCAAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAA 4380

Qy       4851 TACAAGGAATGTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATT 4910
Db       4381 TACAAGGAATGTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATT 4440

Qy       4911 TTCAGTGTACTAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATA 4970
Db       4441 TTCAGTGTACTAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATA 4500

Qy       4971 ATATTCCTGAGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTT 5030
Db       4501 ATATTCCTGAGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTT 4560

Qy       5031 GCGCCCATCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATA 5090
Db       4561 GCGCCCATCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATA 4620

Qy       5091 GAGGATCTTGCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAG 5150
Db       4621 GAGGATCTTGCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAG 4680

Qy       5151 CGAATGATGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAA 5210
Db       4681 CGAATGATGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTCGGTTAA 4740

Qy       5211 AACAAGTGCCTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTG 5270
Db       4741 AACAAGTGCCTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTG 4800

Qy       5271 GGTGTATGAGACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGC 5330
Db       4801 GGTGTATGAGACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGC 4860

Qy       5331 TCTTCAGCCAGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGT 5390
Db       4861 TCTTCAGCCAGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGT 4920

Qy       5391 ATATCAAATCTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCA 5450
Db       4921 ATATCAAATCTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCA 4980

Qy       5451 TACTACAAATCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGT 5510
Db       4981 TACTACAAATCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGT 5040

Qy       5511 TAGCTAAGTGGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGT 5570
Db       5041 TAGCTAAGTGGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGT 5100

Qy       5571 TCCCTGAAAGAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGAT 5630
Db       5101 TCCCTGAAAGAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGAT 5160

Qy       5631 CATCAGGCAGGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTC 5690
Db       5161 CATCAGGCAGGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTC 5220

Qy       5691 TGGTTCTACCACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTAC 5750
Db       5221 TGGTTCTACCACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTAC 5280

Qy       5751 TATTTCCCTGTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTT 5810
Db       5281 TATTTCCCTGTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTT 5340

Qy       5811 CAACAGCCTATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGG 5870
Db       5341 CAACAGCCTATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGG 5400

Qy       5871 TGATGCTTGAAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAG 5930
Db       5401 TGATGCTTGAAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAG 5460

Qy       5931 AAAAGGCATTTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGT 5990
Db       5461 AAAAGGCATTTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGT 5520

Qy       5991 TGGGAGTCGGAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGT 6050
Db       5521 TGGGAGTCGGAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGT 5580

Qy       6051 AAGAAAAAAAAAGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATC 6110
Db       5581 AAGAAAAAAAAAGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATC 5640

Qy       6111 TTAGTAGTCTAGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGA 6170
Db       5641 TTAGTAGTCTAGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGA 5700

Qy       6171 TCAAGTAAATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTT 6230
Db       5701 TCAAGTAAATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTT 5760

Qy       6231 ACAGGTGGATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCA 6290
Db       5761 ACAGGTGGATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCA 5820

Qy       6291 TGTGTTTTGGGGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTG 6350
Db       5821 TGTGTTTTGGGGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTG 5880

Qy       6351 TGTGTTTGAAGTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGAC 6410
Db       5881 TGTGTTTGAAGTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGAC 5940

Qy       6411 ACATTGTCCTATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTT 6470
Db       5941 ACATTGTCCTATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATTTGGGGACAGGTTTT 6000

Qy       6471 CATTTGTCAGATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAA 6530
```


Db 6001 CATTTGTCAGATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAA 6060
Qy 6531 AACTTTGCTAAATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTC 6590
Db 6061 AACTTTGCTAAATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTC 6120
Qy 6591 TGTTGAAGATTAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTA 6650
Db 6121 TGTTGAAGATTAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTA 6180
Qy 6651 AACGGCTTATTTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTC 6710
Db 6181 AACGGCTTATTTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTC 6240
Qy 6711 TAAATTATTCTGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCC 6770
Db 6241 TAAATTATTCTGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCC 6300
Qy 6771 AGATCCTGCCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGA 6830
Db 6301 AGATCCTGCCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGA 6360
Qy 6831 AATGATCTTTCTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCT 6890
Db 6361 AATGATCTTTCTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCT 6420
Qy 6891 GTACCCAGAAAAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTC 6950
Db 6421 GTACCCAGAAAAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTC 6480
Qy 6951 CAACAGCATTTGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATT 7010
Db 6481 CAACAGCATTTGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATT 6540
Qy 7011 TTTTCATGCTTTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAG 7070
Db 6541 TTTTCATGCTTTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAG 6600
Qy 7071 GTACTTCAGCTACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGA 7130
Db 6601 GTACTTCAGCTACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGA 6660
Qy 7131 TGGAGGAAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTT 7190
Db 6661 TGGAGGAAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTT 6720
Qy 7191 CCGTTCAAATTATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATT 7250
Db 6721 CCGTTCAAATTATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATT 6780
Qy 7251 TTTTTAACCATGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCA 7310
Db 6781 TTTTTAACCATGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCA 6840
Qy 7311 GATACTCTGCCAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGT 7370
Db 6841 GATACTCTGCCAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGT 6900
Qy 7371 GGTTATTTTTATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCA 7430
Db 6901 GGTTATTTTTATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCA 6960
Qy 7431 TTTCTTTTCTTTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAG 7490
Db 6961 TTTCTTTTCTTTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAG 7020
Qy 7491 TGCAGTGGTGCAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCT 7550
Db 7021 TGCAGTGGTGCAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCT 7080
Qy 7551 GCCTCAGCCTCCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCT 7610
Db 7081 GCCTCAGCCTCCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCT 7140
Qy 7611 GTATTTTTAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACC 7670
Db 7141 GTATTTTTAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACC 7200
Qy 7671 TCAAGTGATCCGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTAC 7730
Db 7201 TCAAGTGATCCGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTAC 7260
Qy 7731 AGGCGTGAGCCACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCT 7790
Db 7261 AGGCGTGAGCCACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCT 7320
Qy 7791 GACTCTTATCTTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGAT 7850
Db 7321 GACTCTTATCTTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGAT 7380
Qy 7851 TGAGAATGAAATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCA 7910
Db 7381 TGAGAATGAAATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCA 7440
Qy 7911 TGATGTAGGAGGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCC 7970
Db 7441 TGATGTAGGAGGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCC 7500
Qy 7971 TTGGGCAAGTTATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAA 8030
Db 7501 TTGGGCAAGTTATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAA 7560
Qy 8031 TACATATTTCCCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTT 8090
Db 7561 TACATATTTCCCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTT 7620
Qy 8091 TTCTATTGAAGATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAA 8150
Db 7621 TTCTATTGAAGATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAA 7680
Qy 8151 CTGATGAGAAATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTT 8210
Db 7681 CTGATGAGAAATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTT 7740
Qy 8211 TTCTTAATAAATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGT 8270
Db 7741 TTCTTAATAAATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGT 7800
Qy 8271 GCCATTTCTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGC 8330
Db 7801 GCCATTTCTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGC 7860
Qy 8331 CCAGAAGACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGT 8390
Db 7861 CCAGAAGACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGT 7920
Qy 8391 TGCTGCTGAAGTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGC 8450
Db 7921 TGCTGCTGAAGTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGC 7980
Qy 8451 TGACAGTGCGCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGG 8510
Db 7981 TGACAGTGCGCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGG 8040
Qy 8511 CAGTCCACCTGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAG 8570
Db 8041 CAGTCCACCTGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAG 8100
Qy 8571 CGCTGTCCTCTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGA 8630
Db 8101 CGCTGTCCTCTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGA 8160
Qy 8631 CAGTGCTGATAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAG 8690
Db 8161 CAGTGCTGATAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAG 8220

```
Qy       8691 ATCAGCAGTTGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCC 8750
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8221 ATCAGCAGTTGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCC 8280

Qy       8751 TCCTCCGGATGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGA 8810
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8281 TCCTCCGGATGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGA 8340

Qy       8811 GACCGCAGCTGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTT 8870
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8341 GACCGCAGCTGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTT 8400

Qy       8871 ATCCGTGGGGGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAG 8930
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8401 ATCCGTGGGGGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAG 8460

Qy       8931 TATCACCTGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCA 8990
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8461 TATCACCTGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCA 8520

Qy       8991 GAGACTGGAGCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCT 9050
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8521 GAGACTGGAGCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCT 8580

Qy       9051 TAGAAGTCGACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCC 9110
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       8581 TAGAAGTCGACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCTC 8640

Qy       9111 TAGTTTCTGTAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATC 9170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8641 TAGTTTCTGTAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATC 8700

Qy       9171 TGCGAAGGGGAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGA 9230
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8701 TGCGAAGGGGAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGA 8760

Qy       9231 CTCTAGCACTTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGT 9290
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8761 CTCTAGCACTTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGT 8820

Qy       9291 CCAAGGGGCCGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCC 9350
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8821 CCAAGGGGCCGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCC 8880

Qy       9351 TGCACCGGATAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTC 9410
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8881 TGCACCGGATAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTC 8940

Qy       9411 TCAGCCCACCCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGA 9470
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8941 TCAGCCCACCCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGA 9000

Qy       9471 CAGCCCGTGGGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGC 9530
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9001 CAGCCCGTGGGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGC 9060

Qy       9531 AGCAGAATTTGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGT 9590
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9061 AGCAGAATTTGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGT 9120

Qy       9591 TGACCAGCCCGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAG 9650
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9121 TGACCAGCCCGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAG 9180

Qy       9651 AGGCATCACATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAA 9710
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9181 AGGCATCACATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAA 9240

Qy       9711 ACCAAGGTTCCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCC 9770
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9241 ACCAAGGTTCCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCC 9300

Qy       9771 CTATGGTGGCTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTC 9830
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9301 CTATGGTGGCTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTC 9360

Qy       9831 GACGTTAGAAGGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAAT 9890
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9361 GACGTTAGAAGGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAAT 9420

Qy       9891 AGCTAATCCAGGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGC 9950
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9421 AGCTAATCCAGGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGC 9480

Qy       9951 CAGAAGTTCTTGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9481 CAGAAGTTCTTGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 9530
```

RESULT 4
US-09-560-594-3
; Sequence 3, Application US/09560594
; Patent No. 6242590
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowsert
; TITLE OF INVENTION: ANTISENSE MODULATION OF ZINC FINGER PROTEIN-217 EXPRESSION
; FILE REFERENCE: RTS-0144
; CURRENT APPLICATION NUMBER: US/09/560,594
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (272)...(3418)
US-09-560-594-3

Query Match 16.4%; Score 1639.4; DB 3; Length 5632;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 2533
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 60

Qy       2534 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 2593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 120

Qy       2594 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 2653
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 180

Qy       2654 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 2713
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 240

Qy       2714 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300

Qy       2774 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 2833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 360

Qy       2834 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 2893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 420

Qy       2894 TACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 2953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 480
```

```
Qy       2954 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 540

Qy       3014 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600

Qy       3074 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 660

Qy       3134 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 3193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 720

Qy       3194 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 3253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 780

Qy       3254 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 3313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 840

Qy       3314 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 900

Qy       3374 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 960

Qy       3434 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 1020

Qy       3494 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 1080

Qy       3554 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 3613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 1140

Qy       3614 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 3673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 1200

Qy       3674 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 3733
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 1260

Qy       3734 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 3793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 1320

Qy       3794 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 3853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 1380

Qy       3854 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 3913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 1440

Qy       3914 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 3973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 1500

Qy       3974 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 4033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 1560

Qy       4034 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 4093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 1620

Qy       4094 CGAAGGAATCCATCTGGGTAA 4114
              |||||||||||||||| |||
Db       1621 CGAAGGAATCCATCTGGATAA 1641
```

RESULT 5
US-09-949-016-862
; Sequence 862, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 862
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-862

Query Match 16.4%; Score 1639.4; DB 4; Length 5632;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 2533
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 60

Qy       2534 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 2593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 120

Qy       2594 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 2653
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 180

Qy       2654 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 2713
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 240

Qy       2714 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300

Qy       2774 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 2833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 360

Qy       2834 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 2893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 420

Qy       2894 TACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 2953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 480

Qy       2954 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 540
```

```
Qy       3014 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600

Qy       3074 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 660

Qy       3134 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 3193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 720

Qy       3194 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 3253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 780

Qy       3254 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 3313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 840

Qy       3314 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 900

Qy       3374 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 960

Qy       3434 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 1020

Qy       3494 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 1080

Qy       3554 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 3613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 1140

Qy       3614 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 3673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 1200

Qy       3674 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 3733
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 1260

Qy       3734 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 3793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 1320

Qy       3794 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 3853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 1380

Qy       3854 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 3913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 1440

Qy       3914 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 3973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 1500

Qy       3974 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCCTCTGGATGA 4033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCCTCTGGATGA 1560

Qy       4034 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 4093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 1620

Qy       4094 CGAAGGAATCCATCTGGGTAA 4114
                ||||||||||||||||| |||
Db       1621 CGAAGGAATCCATCTGGATAA 1641
```

RESULT 6
US-09-949-016-4262
; Sequence 4262, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4262
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4262

Query Match 16.4%; Score 1639.4; DB 4; Length 5632;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 2533
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 60

Qy       2534 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 2593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 120

Qy       2594 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 2653
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 180

Qy       2654 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 2713
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 240

Qy       2714 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300

Qy       2774 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 2833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 360

Qy       2834 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 2893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 420

Qy       2894 TACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 2953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 480

Qy       2954 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 540

Qy       3014 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600
```

```
Qy       3074 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 660

Qy       3134 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 3193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 720

Qy       3194 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 3253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 780

Qy       3254 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 3313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 840

Qy       3314 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 900

Qy       3374 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 960

Qy       3434 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 1020

Qy       3494 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 1080

Qy       3554 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 3613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 1140

Qy       3614 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 3673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 1200

Qy       3674 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 3733
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 1260

Qy       3734 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 3793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 1320

Qy       3794 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 3853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 1380

Qy       3854 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 3913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 1440

Qy       3914 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 3973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 1500

Qy       3974 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 4033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 1560

Qy       4034 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 4093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 1620

Qy       4094 CGAAGGAATCCATCTGGGTAA 4114
             |||||||||||||||| |||
Db       1621 CGAAGGAATCCATCTGGATAA 1641
```

```
RESULT 7
US-08-892-695-10
; Sequence 10, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
;  APPLICANT: Gray, Joe W
;  APPLICANT: Collins, Collin
;  APPLICANT: Hwang, Soo In
;  APPLICANT: Godfrey, Tony
;  APPLICANT: Kowel, David
;  APPLICANT: Rommens, Johanna
;  TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
;  FILE REFERENCE: 2500.124US3
;  CURRENT APPLICATION NUMBER: US/08/892,695A
;  CURRENT FILING DATE: 1997-07-15
;  EARLIER APPLICATION NUMBER: 08/785,532
;  EARLIER FILING DATE: 1997-01-17
;  EARLIER APPLICATION NUMBER: 08/731,499
;  EARLIER FILING DATE: 1996-10-16
;  EARLIER APPLICATION NUMBER: 08/680,395
;  EARLIER FILING DATE: 1996-07-15
;  NUMBER OF SEQ ID NOS: 59
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 3186
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:ZABC1 Open
;   OTHER INFORMATION: Reading Frame
US-08-892-695-10

  Query Match             15.6%;  Score 1559.6;  DB 4;  Length 3186;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1565;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;
```

```
Qy       8286 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACATCTCT 8345
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1482 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACATCTCT 1541

Qy       8346 GAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 8405
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1542 GAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 1601

Qy       8406 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCGCAAAC 8465
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1602 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCGCAAAC 1661

Qy       8466 CAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCTGCAAA 8525
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1662 CAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCTGCAAA 1721

Qy       8526 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTCTCACC 8585
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1722 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTCTCACC 1781

Qy       8586 AGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGATAAAGT 8645
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1782 AGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGATAAAGT 1841

Qy       8646 GAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAC 8705
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1842 GAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAC 1901

Qy       8706 TCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGATGGCAG 8765
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1902 TCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGATGGCAG 1961

Qy       8766 TACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCTGACTG 8825
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1962 TACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCTGACTG 2021

Qy       8826 CAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCT 8885
```

```
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2022 CAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCT 2081

Qy       8886 TCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGTCCATT 8945
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2082 TCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGTCCATT 2141

Qy       8946 TTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAGCATAA 9005
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2142 TTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAGCATAA 2201

Qy       9006 ATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGACGTAC 9065
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2202 ATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGACGTAC 2261

Qy       9066 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGTAAACC 9125
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2262 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCTCTAGTTTCTGTAAACC 2321

Qy       9126 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 9185
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2322 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 2381

Qy       9186 GAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 9245
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2382 GAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 2441

Qy       9246 CCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCCGCCAC 9305
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2442 CCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCCGCCAC 2501

Qy       9306 CAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGATAAGAC 9365
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2502 CAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGATAAGAC 2561

Qy       9366 AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACCCTCGG 9425
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2562 AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACCCTCGG 2621

Qy       9426 CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGGGCACC 9485
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2622 CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGGGCACC 2681

Qy       9486 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTTGGTGA 9545
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2682 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTTGGTGA 2741

Qy       9546 GCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCCGGGGC 9605
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2742 GCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCCGGGGC 2801

Qy       9606 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT 9665
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2802 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT 2861

Qy       9666 GTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTCCTGAG 9725
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2862 GTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTCCTGAG 2921

Qy       9726 CTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGCTCCGG 9785
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2922 CTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGCTCCGG 2981

Qy       9786 GCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTAT 9845
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2982 GCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTCT 3041

Qy       9846 TGCATGAGGGGCGT 9859
                  ||  || |  ||
Db       3042 TGGTGGATGTCAGT 3055

RESULT 8
```

```
US-08-680-395-3
; Sequence 3, Application US/08680395
; Patent No. 5892010
;  GENERAL INFORMATION:
;    APPLICANT:  Gray, Joe W.
;    APPLICANT:  Collins, Colin
;    APPLICANT:  Hwang, Soo-in
;    APPLICANT:  Godfrey, Tony
;    APPLICANT:  Kowbel, David
;    APPLICANT:  Rommens, Johanna
;    TITLE OF INVENTION:  Genes from the 20q13 Amplicon and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  40
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/680,395
;      FILING DATE:  15-JUL-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  02307O-068900US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1507 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1..1507
;      OTHER INFORMATION:  /note= "cDNA clone cc49 of 6-7kb
;      OTHER INFORMATION:  transcript with homology to C2H2 zinc
;      OTHER INFORMATION:  finger genes"
US-08-680-395-3

  Query Match           13.9%;  Score 1387.2;  DB 2;  Length 1507;
  Best Local Similarity  97.3%;  Pred. No. 0;
  Matches 1470;  Conservative    3;  Mismatches   32;  Indels    6;  Gaps    6;

Qy       2424 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGA 2483
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGC 61

Qy       2484 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGT 2543
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGT 121

Qy       2544 TTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGG 2603
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGG 181

Qy       2604 ACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTTGGAACCCAGG 2663
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGG-TTTTGGAACCCAGG 240

Qy       2664 GCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTCCAGAAATTGC 2723
                  |||||||| |||||||||||| ||||||| |||||||||| ||||||||||||||||||
```

```
Db        241 ACTGTTAGA-GGGTGAGTGACAGGTCTTAC-AGTGGCCTTAATCCAACTCCAGAAATTGC 298

Qy       2724 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 2783

Db        299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 358

Qy       2784 CTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAG 2843

Db        359 CTCTTAATGTACATGGAT-GGCCAAGAGTGATTGGCAGCTCTCTTGCCAGTCCGATGGAG 417

Qy       2844 ATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGAA 2903

Db        418 ATGGA-GATGCCTTGTCAATGAAAGGGCCCNCTGTTGTCAATTCCGAGCTACACAAAGAA 476

Qy       2904 AAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 2962

Db        477 AAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 536

Qy       2963 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 3022

Db        537 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 596

Qy       3023 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 3082

Db        597 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 656

Qy       3083 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 3142

Db        657 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 716

Qy       3143 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 3202

Db        717 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 776

Qy       3203 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 3262

Db        777 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 836

Qy       3263 TCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 3322

Db        837 TCACATGCGGACRCATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 896

Qy       3323 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTC 3382

Db        897 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTC 956

Qy       3383 TCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 3442

Db        957 TCCTTGCAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 1016

Qy       3443 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTC 3502

Db       1017 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTC 1076

Qy       3503 TCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGACC 3562

Db       1077 TCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGACC 1136

Qy       3563 AAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGTT 3622

Db       1137 AAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGTT 1196

Qy       3623 CACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTGCCAAGAAGT 3682

Db       1197 CACCACCTTCCAGGCTTGGCAKCTGGCTACCAAAGGAAWAGTTGCCATTTGCCAAGAAGT 1256

Qy       3683 GAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGAGAAGGAGCT 3742

Db       1257 GAAGGAATTGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGAGAAGGAGCT 1316

Qy       3743 TGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACA 3802

Db       1317 TGGAGAAACAAATAAGAACCATTGTGCAGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACA 1376

Qy       3803 CTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAA 3862

Db       1377 CTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAA 1436

Qy       3863 GCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCTGGTCTTGCA 3922

Db       1437 GCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCTGGTCTTGCA 1496

Qy       3923 CTCCAGGGTCC 3933

Db       1497 CTCCAGGGTCC 1507
```

RESULT 9
US-08-892-695-3
; Sequence 3, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892,695A
; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cc49
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (447)
; OTHER INFORMATION: N is A, G, G, T, or U
US-08-892-695-3

Query Match 13.9%; Score 1387.2; DB 4; Length 1507;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

```
Qy       2424 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGA 2483

Db          2 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGC 61

Qy       2484 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGT 2543

Db         62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGT 121

Qy       2544 TTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGG 2603

Db        122 TTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGG 181

Qy       2604 ACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTTGGAACCCAGG 2663

Db        182 ACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGG-TTTTGGAACCCAGG 240

Qy       2664 GCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTCCAGAAATTGC 2723

Db        241 ACTGTTAGA-GGGTGAGTGACAGGTCTTAC-AGTGGCCTTAATCCAACTCCAGAAATTGC 298
```

```
Qy       2724 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 2783
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 358

Qy       2784 CTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAG 2843
             |||||||||||||||||| |||||  |||||||||||||||||||| ||||||||||||||
Db        359 CTCTTAATGTACATGGAT-GGCCAAGAGTGATTGGCAGCTCTCTTGCCAGTCCGATGGAG 417

Qy       2844 ATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGAA 2903
             ||||| |||||||||||||||||||| || |||||||   | |      |||||
Db        418 ATGGA-GATGCCTTGTCAATGAAAGGGCCCNCTGTTGTCAATTCCGAGCTACACAAAGAA 476

Qy       2904 AAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 2962
             |||||||| |||| |||||||||| ||||||||||||||||||||||||||||||||||||
Db        477 AAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 536

Qy       2963 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 3022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 596

Qy       3023 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 3082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 656

Qy       3083 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 3142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 716

Qy       3143 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 3202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 776

Qy       3203 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 3262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 836

Qy       3263 TCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 3322
             ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCACATGCGGACRCATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 896

Qy       3323 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTC 3382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTC 956

Qy       3383 TCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 3442
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 TCCTTGCAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 1016

Qy       3443 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTC 3502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTC 1076

Qy       3503 TCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGACC 3562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 TCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGACC 1136

Qy       3563 AAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGTT 3622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1137 AAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGTT 1196

Qy       3623 CACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTGCCAAGAAGT 3682
             ||||||||||||||||||||:|||||||||||||||||:||||||||||||||||||||||
Db       1197 CACCACCTTCCAGGCTTGGCAKCTGGCTACCAAAGGAAWAGTTGCCATTTGCCAAGAAGT 1256

Qy       3683 GAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGAGAAGGAGCT 3742
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1257 GAAGGAATTGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGAGAAGGAGCT 1316

Qy       3743 TGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACA 3802
             |||||||||||||||  | ||||||||||||||||||||||||||||||||||||||||
Db       1317 TGGAGAAACAAATAAGAACCATTGTGCAGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACA 1376

Qy       3803 CTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAA 3862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1377 CTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAA 1436

Qy       3863 GCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCTGGTCTTGCA 3922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1437 GCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCTGGTCTTGCA 1496

Qy       3923 CTCCAGGGTCC 3933
             ||||||||||| 
Db       1497 CTCCAGGGTCC 1507
```

RESULT 10
US-09-949-016-34699/c
; Sequence 34699, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34699
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-34699

Query Match 6.0%; Score 600.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 8.1e-141;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2843 GATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGA 2902
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGA 542

Qy       2903 AAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 2962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 482

Qy       2963 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 3022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 422

Qy       3023 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 3082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 362

Qy       3083 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 3142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 302

Qy       3143 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 3202
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 242

Qy       3203 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 3262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 182

Qy       3263 TCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 3322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 122
```

```
Qy       3323 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTC 3382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTC 62

Qy       3383 TCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 3442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 2

Qy       3443 G 3443
              |
Db          1 G 1


RESULT 11
US-09-949-016-34700/c
; Sequence 34700, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34700
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-34700

  Query Match             6.0%;  Score 600.6;  DB 4;  Length 601;
  Best Local Similarity   99.8%;  Pred. No. 8.1e-141;
  Matches  600;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3470 TGCTTTCGGTACCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAG 3529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGCTTTCGGTACCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAG 542

Qy       3530 GGAGGACTTCCTGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAA 3589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAGGACTTCCTGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAA 482

Qy       3590 GCCTGTCAGATGCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGC 3649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCTGTCAGATGCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGC 422

Qy       3650 TACCAAAGGAAAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCAC 3709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACCAAAGGAAAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCAC 362

Qy       3710 CGACAACGACGATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGC 3769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGACAACGACGATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGC 302

Qy       3770 AGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGA 3829
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 WGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGA 242

Qy       3830 CGCGGATCCCAAGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAA 3889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGCGGATCCCAAGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAA 182

Qy       3890 AGCTTTCAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAG 3949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCTTTCAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAG 122

Qy       3950 GGCCGGCGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCC 4009
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCCGGCGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCC 62

Qy       4010 TGACCTCGCCGCCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGA 4069
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TGACCTCGCCGCCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGA 2

Qy       4070 A 4070
              |
Db          1 A 1


RESULT 12
US-09-949-016-34701/c
; Sequence 34701, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34701
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-34701

  Query Match             6.0%;  Score 600.6;  DB 4;  Length 601;
  Best Local Similarity   99.8%;  Pred. No. 8.1e-141;
  Matches  600;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5250 CTTGAATTCTTCGGAAAACTGGGTGTATGAGACCCACGATGAATTTGCCCACACGATTGA 5309
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTGAATTCTTCGGAAAACTGGGTGTATGAGACCCACGATGAATTTGCCCACACGATTGA 542

Qy       5310 TTGGACTCTTCCTTCACCTGCTCTTCAGCCAGTGCCAGTTCCTTTTCTGATCATGTGATT 5369
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGGACTCTTCCTTCACCTGCTCTTCAGCCAGTGCCAGTTCCTTTTCTGATCATGTGATT 482

Qy       5370 GACGTGAGAACTGTAGTCTGTATATCAAATCTTTAGAATGTTTTTGAGTTTCCTGGGACA 5429
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GACGTGAGAACTGTAGTCTGTATATCAAATCTTTAGAATGTTTTTGAGTTTCCTGGGACA 422

Qy       5430 CAGGAAACCCAGCACTTAGCATACTACAAATCTAATGTCTTAATGGCATCATAAAAAGAG 5489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGGAAACCCAGCACTTAGCATACTACAAATCTAATGTCTTAATGGCATCATAAAAAGAG 362

Qy       5490 GCTTTAAACACAGACTCCAGTTAGCTAAGTGGTTTCTGCTAGTGCCGGTACTGTTGCAGG 5549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTTTAAACACAGACTCCAGTTAGCTAAGTGGTTTCTGCTAGTGCCGGTACTGTTGCAGG 302

Qy       5550 GGCCCTGTGAGATGCCCCAGTTCCCTGAAAGAAATGAAAAGGCCAGTTACCGGTAGGTGG 5609
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 RGCCCTGTGAGATGCCCCAGTTCCCTGAAAGAAATGAAAAGGCCAGTTACCGGTAGGTGG 242

Qy       5610 TGTGGAAAACATGGGCTAGATCATCAGGCAGGACAGAATGCCTGGCTGTGGGTGGGAGCA 5669
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGGAAAACATGGGCTAGATCATCAGGCAGGACAGAATGCCTGGCTGTGGGTGGGAGCA 182

Qy       5670 CCCCAGCTTGGCGTTGAGTTCTGGTTCTACCACTGCGTTGTTTTGTGACCAATTATGAGT 5729
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 CCCCAGCTTGGCGTTGAGTTCTGGTTCTACCACTGCGTTGTTTTGTGACCAATTATGAGT 122

Qy       5730 TGCTTAACCTTTCTTTGCTACTATTTCCCTGTTTGCAAAATGGTTCATTGACCCCTGTCT 5789
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCTTAACCTTTCTTTGCTACTATTTCCCTGTTTGCAAAATGGTTCATTGACCCCTGTCT 62

Qy       5790 TCCACCTCCCAAGGACAATTTCAACAGCCTATTTGTAAAAAGATCACAGTCCTTTAAAAA 5849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCCACCTCCCAAGGACAATTTCAACAGCCTATTTGTAAAAAGATCACAGTCCTTTAAAAA 2

Qy       5850 A 5850
              |
Db          1 A 1
```

RESULT 13
US-09-949-016-34702/c
; Sequence 34702, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34702
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-34702

Query Match 6.0%; Score 600.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 8.1e-141;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       7580 AGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTTTAGTAGAGATGGGGTTTCACCA 7639
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTTTAGTAGAGATGGGGTTTCACCA 542

Qy       7640 TGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATCCGCTCACCTTGGCCTCCCA 7699
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATCCGCTCACCTTGGCCTCCCA 482

Qy       7700 TAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGACAAAGT 7759
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGACAAAGT 422

Qy       7760 TCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTTATCTTATTAAAAGCTACAGTAT 7819
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTTATCTTATTAAAAGCTACAGTAT 362

Qy       7820 TTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAATGAAATGAGAATCTATTTAGTAG 7879
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAATGAAATGAGAATCTATTTAGTAG 302

Qy       7880 TCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTAGGAGGCTGCGTAGATTTGAAAT 7939
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTAGGAGGCTGCGTAGATTTGAAAT 242

Qy       7940 TTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCAAGTTATTTAACCTTTTTGTGCT 7999
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCAAGTTATTTAACCTTTTTGTGCT 182

Qy       8000 TTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATATTTCCCTAGGGCTGTTAGGAAGA 8059
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATATTTCCCTAGGGCTGTTAGGAAGA 122

Qy       8060 TTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATTGAAGATAGGCATTCATAATTTC 8119
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATTGAAGATAGGCATTCATAATTTC 62

Qy       8120 AAATATTCATTACAGTAAGGATGATAAAGAACTGATGAGAAATCCTATGTGATAGTAGAT 8179
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAATATTCATTACAGTAAGGATGATAAAGAACTGATGAGAAATCCTATGTGATAGTAGAT 2

Qy       8180 C 8180
              |
Db          1 C 1
```

RESULT 14
US-09-949-016-151879/c
; Sequence 151879, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 151879
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-151879

Query Match 6.0%; Score 600.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 8.1e-141;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2843 GATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGA 2902
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGA 542

Qy       2903 AAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 2962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 482

Qy       2963 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 3022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 422

Qy       3023 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 3082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 362

Qy       3083 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 3142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 302

Qy       3143 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 3202
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 242

Qy       3203 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 3262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 182
```

```
Qy       3263 TCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 3322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 122

Qy       3323 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTC 3382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTC 62

Qy       3383 TCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 3442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 2

Qy       3443 G 3443
              |
Db          1 G 1


RESULT 15
US-09-949-016-151880/c
; Sequence 151880, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 151880
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-151880

  Query Match             6.0%;  Score 600.6;  DB 4;  Length 601;
  Best Local Similarity   99.8%;  Pred. No. 8.1e-141;
  Matches  600;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3470 TGCTTTCGGTACCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAG 3529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 TGCTTTCGGTACCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAG 542

Qy       3530 GGAGGACTTCCTGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAA 3589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GGAGGACTTCCTGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAA 482

Qy       3590 GCCTGTCAGATGCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGC 3649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 GCCTGTCAGATGCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGC 422

Qy       3650 TACCAAAGGAAAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCAC 3709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 TACCAAAGGAAAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCAC 362

Qy       3710 CGACAACGACGATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGC 3769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CGACAACGACGATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGC 302

Qy       3770 AGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGA 3829
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 WGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGA 242

Qy       3830 CGCGGATCCCAAGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAA 3889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CGCGGATCCCAAGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAA 182

Qy       3890 AGCTTTCAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAG 3949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AGCTTTCAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAG 122

Qy       3950 GGCCGGCGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCC 4009
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GGCCGGCGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCC 62

Qy       4010 TGACCTCGCCGCCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGA 4069
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TGACCTCGCCGCCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGA 2

Qy       4070 A 4070
              |
Db          1 A 1


Search completed: August  1, 2005, 01:55:37
Job time : 1034 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 09:38:42 ; Search time 3126 Seconds
(without alignments)
18937.112 Million cell updates/sec

Title: US-08-731-499-9_COPY_1_10000
Perfect score: 10000
Sequence: 1 CCATCATATTTCTTATTTTT.....ATTCTTAGATACGGCAGTGG 10000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

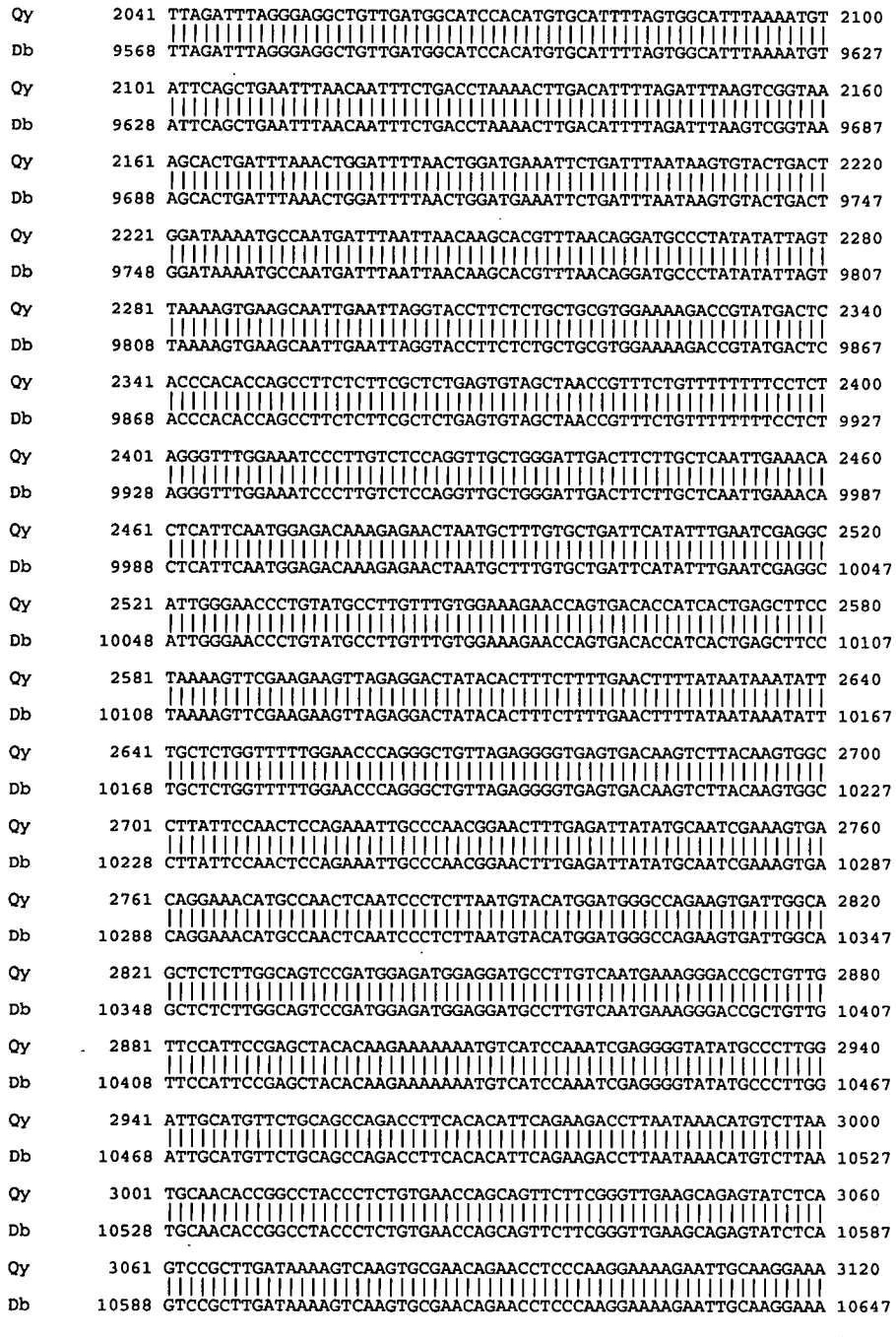
SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	9980.6	99.8	36022	11	ACN44986	Acn44986 Human gen
	2	9651.6	96.5	10282	2	AAV09023	Aav09023 Homo sapi
	3	6382.2	63.8	14906	4	AAK81093	Aak81093 Human imm
	4	1639.4	16.4	5632	8	ACC50334	Acc50334 Breast ca
	5	1639.4	16.4	5632	10	ADD14635	Add14635 Human src
	6	1639.4	16.4	5632	13	ADR66761	Adr66761 Human pro
	7	1639.4	16.4	5632	13	ADR65858	Adr65858 Human pro
	8	1639.4	16.4	5653	12	ADP07267	Adp07267 Human ZNF
	9	1633.4	16.3	1797	4	AAK61822	Aak61822 Human imm
	10	1559.6	15.6	5252	11	ACN44987	Acn44987 Human mRN
	11	1554.8	15.5	3183	2	AAV09024	Aav09024 Homo sapi
	12	1387.2	13.9	1507	2	AAV04698	Aav04698 Homo sapi
c	13	1006.8	10.1	124987	12	ADP03055	Adp03055 Human hou
c	14	1006.8	10.1	124990	13	ADS88553	Ads88553 Human hou
	15	820	8.2	26345	11	ACN44984	Acn44984 Mouse gen
	16	726.4	7.3	3016	11	ACN44985	Acn44985 Mouse mRN
	17	566	5.7	566	4	AAK81095	Aak81095 Human imm
c	18	530	5.3	530	12	ACH76516	Ach76516 Human gen
	19	335.8	3.4	469	2	AAV88042	Aav88042 EST clone
c	20	256.8	2.6	501	12	ACH67097	Ach67097 Human gen
c	21	251	2.5	251	12	ACH90216	Ach90216 Human gen
	22	233.8	2.3	37314	4	AAK71358	Aak71358 Human imm
c	23	233.8	2.3	172570	6	ABQ88207	Abq88207 Human ost
	24	227.2	2.3	104245	12	ADG86768	Adg86768 Human clo
	25	227.2	2.3	104245	12	ADL34706	Adl34706 Human PPA
c	26	227.2	2.3	170245	12	ADP13586	Adp13586 Renal cel
c	27	225	2.2	25541	4	AAK76168	Aak76168 Human imm
	28	225	2.2	32248	4	AAS28368	Aas28368 Genomic s
	29	225	2.2	32248	10	ADG41564	Adg41564 Human res
	30	225	2.2	32248	11	ADI97338	Adi97338 Human res
c	31	225	2.2	34435	4	AAK76172	Aak76172 Human imm
	32	224.4	2.2	7815	13	ADS34410	Ads34410 POSH prot
	33	224.4	2.2	78925	3	AAC89888	Aac89888 Human FN
	34	224.2	2.2	298	3	AAC07753	Aac07753 Human sec
	35	224	2.2	66804	6	ABK87050	Abk87050 Human tra
	36	224	2.2	66804	10	ADG88330	Adg88330 Human tra
	37	223.6	2.2	3116	4	AAK90031	Aak90031 Human dig
	38	223.6	2.2	3116	5	AAS39688	Aas39688 Genomic s
	39	223.6	2.2	3116	9	ADB32648	Adb32648 Human nov
	40	223.6	2.2	32221	4	ABK42519	Abk42519 Genomic s
	41	223.6	2.2	32221	9	ADB60675	Adb60675 Connectiv
	42	223.6	2.2	96594	10	ADC85476	Adc85476 Human Msf
	43	223.6	2.2	96595	9	ADA02996	Ada02996 Human Msf
	44	223.6	2.2	96595	10	ADB72734	Adb72734 Human Msf
	45	223.6	2.2	96595	12	ADM74591	Adm74591 Human car

ALIGNMENTS

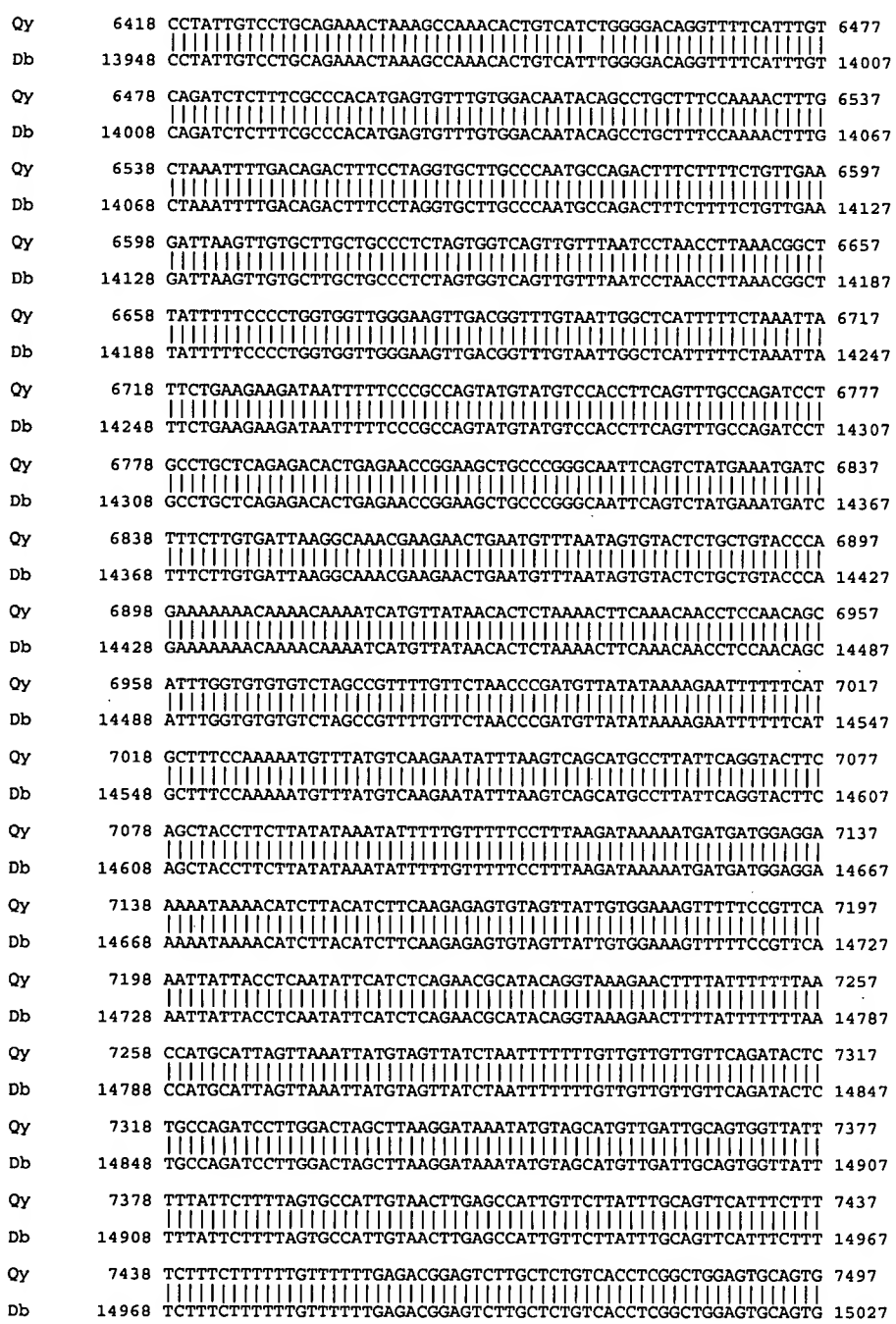
RESULT 1
ACN44986
ID ACN44986 standard; DNA; 36022 BP.
XX
AC ACN44986;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genomic sequence hCG37127.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 1708; 0pp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the

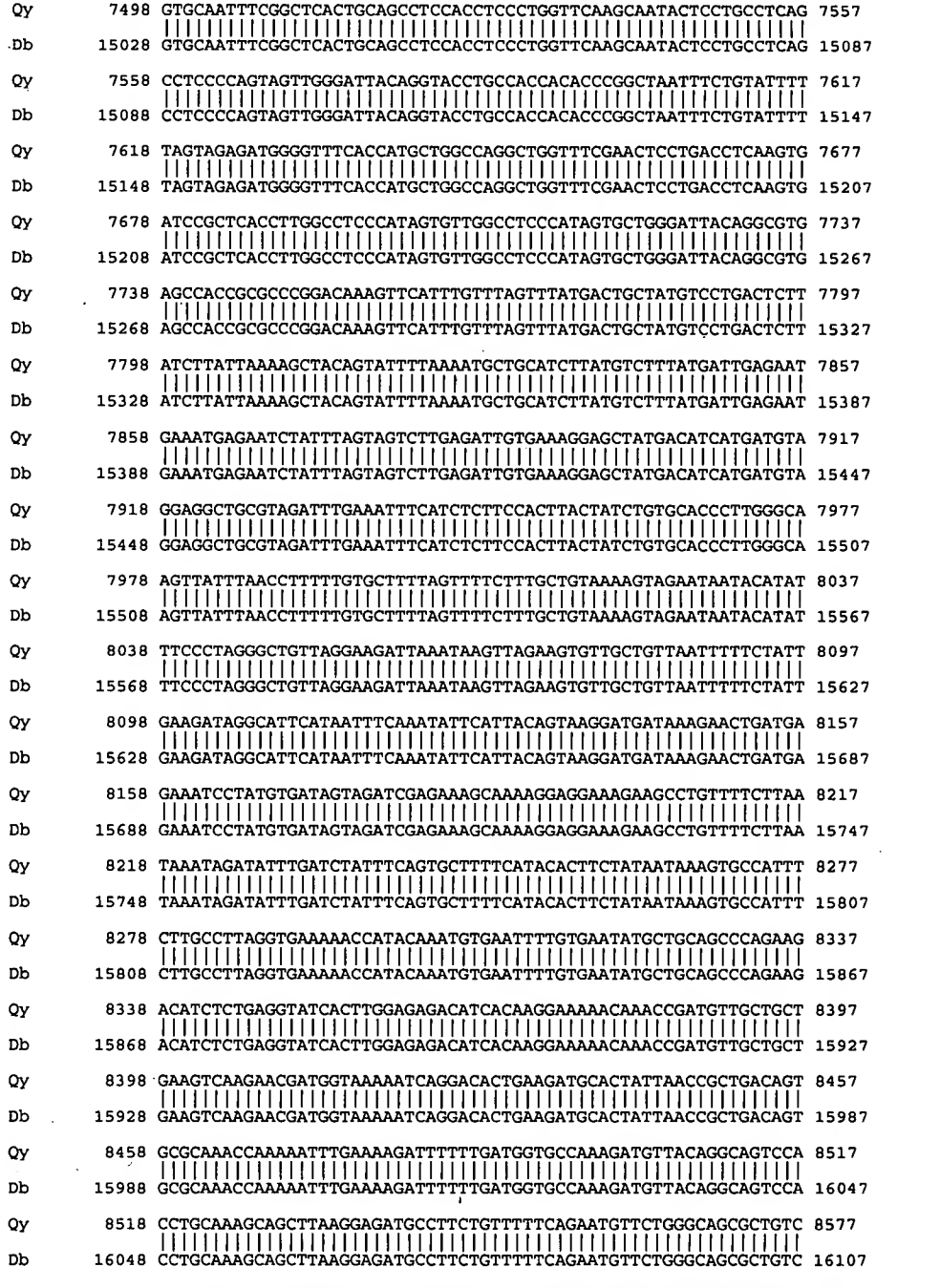
```
CC   CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC   carcinoma including lymphoma. The present sequence is one such CA coding
CC   sequence. Note: This patent is an equivalent to basic patent
CC   US2002182586A1, for which no sequence data was published
XX
SQ   Sequence 36022 BP; 9645 A; 7326 C; 8145 G; 10906 T; 0 U; 0 Other;

  Query Match             99.8%;  Score 9980.6;  DB 11;  Length 36022;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 9996;  Conservative    0;  Mismatches    4;  Indels    3;  Gaps    1;

Qy          1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7528 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 7587

Qy         61 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7588 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 7647

Qy        121 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7648 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 7707

Qy        181 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7708 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 7767

Qy        241 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7768 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 7827

Qy        301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7828 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 7887

Qy        361 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7888 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 7947

Qy        421 AATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTTCAAAT 480
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       7948 AATGTTGAGAGCTGACATAGGGACTAACTTAGTTTTTTTTTTTTTTTTTTTTTTTCAAAT 8007

Qy        481 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAAAACTCCAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8008 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAAAACTCCAC 8067

Qy        541 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8068 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 8127

Qy        601 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8128 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 8187

Qy        661 TTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8188 TTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 8247

Qy        721 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8248 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 8307

Qy        781 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8308 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 8367

Qy        841 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8368 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 8427

Qy        901 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       8428 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 8487

Qy         961 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8488 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 8547

Qy        1021 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 1080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8548 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 8607

Qy        1081 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 1140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8608 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 8667

Qy        1141 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 1200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8668 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 8727

Qy        1201 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 1260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8728 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 8787

Qy        1261 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 1320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8788 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 8847

Qy        1321 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 1380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8848 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 8907

Qy        1381 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 1440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8908 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 8967

Qy        1441 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 1500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8968 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 9027

Qy        1501 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATGA 1560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9028 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATGA 9087

Qy        1561 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 1620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9088 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 9147

Qy        1621 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 1680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9148 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 9207

Qy        1681 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 1740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9208 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 9267

Qy        1741 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 1800
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9268 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 9327

Qy        1801 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9328 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 9387

Qy        1861 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 1920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9388 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 9447

Qy        1921 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 1980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9448 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 9507

Qy        1981 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 2040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9508 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 9567
```


Qy 2041 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 2100
Db 9568 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 9627
Qy 2101 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 2160
Db 9628 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 9687
Qy 2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 2220
Db 9688 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 9747
Qy 2221 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 2280
Db 9748 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 9807
Qy 2281 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 2340
Db 9808 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 9867
Qy 2341 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTCT 2400
Db 9868 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTCT 9927
Qy 2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 2460
Db 9928 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 9987
Qy 2461 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520
Db 9988 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 10047
Qy 2521 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2580
Db 10048 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 10107
Qy 2581 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 2640
Db 10108 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 10167
Qy 2641 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2700
Db 10168 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 10227
Qy 2701 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2760
Db 10228 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 10287
Qy 2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2820
Db 10288 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 10347
Qy 2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2880
Db 10348 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 10407
Qy 2881 TTCCATTCCGAGCTACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 2940
Db 10408 TTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 10467
Qy 2941 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 3000
Db 10468 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 10527
Qy 3001 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 3060
Db 10528 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 10587
Qy 3061 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 3120
Db 10588 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 10647
Qy 3121 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 3180
Db 10648 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 10707
Qy 3181 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 3240
Db 10708 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 10767
Qy 3241 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 3300
Db 10768 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 10827
Qy 3301 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 3360
Db 10828 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 10887
Qy 3361 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 3420
Db 10888 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 10947
Qy 3421 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 3480
Db 10948 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 11007
Qy 3481 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 3540
Db 11008 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 11067
Qy 3541 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 3600
Db 11068 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 11127
Qy 3601 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 3660
Db 11128 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 11187
Qy 3661 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 3720
Db 11188 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 11247
Qy 3721 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 3780
Db 11248 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 11307
Qy 3781 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 3840
Db 11308 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 11367
Qy 3841 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 3900
Db 11368 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 11427
Qy 3901 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 3960
Db 11428 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 11487
Qy 3961 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 4020
Db 11488 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 11547
Qy 4021 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 4080
Db 11548 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 11607
Qy 4081 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 4140
Db 11608 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 11667
Qy 4141 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 4200
Db 11668 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 11727
Qy 4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 4260

```
Db      11728 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 11787
Qy       4261 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 4320
Db      11788 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 11847
Qy       4321 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 4380
Db      11848 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 11907
Qy       4381 TCACAAAACACACTTTAAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 4440
Db      11908 TCACAAAACACACTTTAAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 11967
Qy       4441 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 4500
Db      11968 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 12027
Qy       4501 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 4560
Db      12028 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 12087
Qy       4561 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 4620
Db      12088 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 12147
Qy       4621 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 4680
Db      12148 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 12207
Qy       4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAG---TGATCAACTAGAGC 4737
Db      12208 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATGATCAACTAGAGC 12267
Qy       4738 CAGGAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGC 4797
Db      12268 CAGGAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGC 12327
Qy       4798 AAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGG 4857
Db      12328 AAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGG 12387
Qy       4858 AATGTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTG 4917
Db      12388 AATGTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTG 12447
Qy       4918 TACTAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCC 4977
Db      12448 TACTAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCC 12507
Qy       4978 TGAGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCA 5037
Db      12508 TGAGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCA 12567
Qy       5038 TCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATC 5097
Db      12568 TCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATC 12627
Qy       5098 TTGCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGA 5157
Db      12628 TTGCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGA 12687
Qy       5158 TGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGT 5217
Db      12688 TGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTCGGTTAAAACAAGT 12747
Qy       5218 GCCTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTAT 5277
Db      12748 GCCTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTAT 12807
Qy       5278 GAGACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAG 5337
```

```
Db      12808 GAGACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAG 12867
Qy       5338 CCAGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAA 5397
Db      12868 CCAGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAA 12927
Qy       5398 ATCTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACA 5457
Db      12928 ATCTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACA 12987
Qy       5458 AATCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAA 5517
Db      12988 AATCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAA 13047
Qy       5518 GTGGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGA 5577
Db      13048 GTGGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGA 13107
Qy       5578 AAGAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGG 5637
Db      13108 AAGAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGG 13167
Qy       5638 CAGGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCT 5697
Db      13168 CAGGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCT 13227
Qy       5698 ACCACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCC 5757
Db      13228 ACCACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCC 13287
Qy       5758 CTGTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGC 5817
Db      13288 CTGTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGC 13347
Qy       5818 CTATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCT 5877
Db      13348 CTATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCT 13407
Qy       5878 TGAAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGC 5937
Db      13408 TGAAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGC 13467
Qy       5938 ATTTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGT 5997
Db      13468 ATTTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGT 13527
Qy       5998 CGGAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAA 6057
Db      13528 CGGAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAA 13587
Qy       6058 AAAAGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAG 6117
Db      13588 AAAAGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAG 13647
Qy       6118 TCTAGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTA 6177
Db      13648 TCTAGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTA 13707
Qy       6178 AATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTG 6237
Db      13708 AATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTG 13767
Qy       6238 GATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTT 6297
Db      13768 GATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTT 13827
Qy       6298 TGGGGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTT 6357
Db      13828 TGGGGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTT 13887
Qy       6358 GAAGTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGT 6417
Db      13888 GAAGTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGT 13947
```

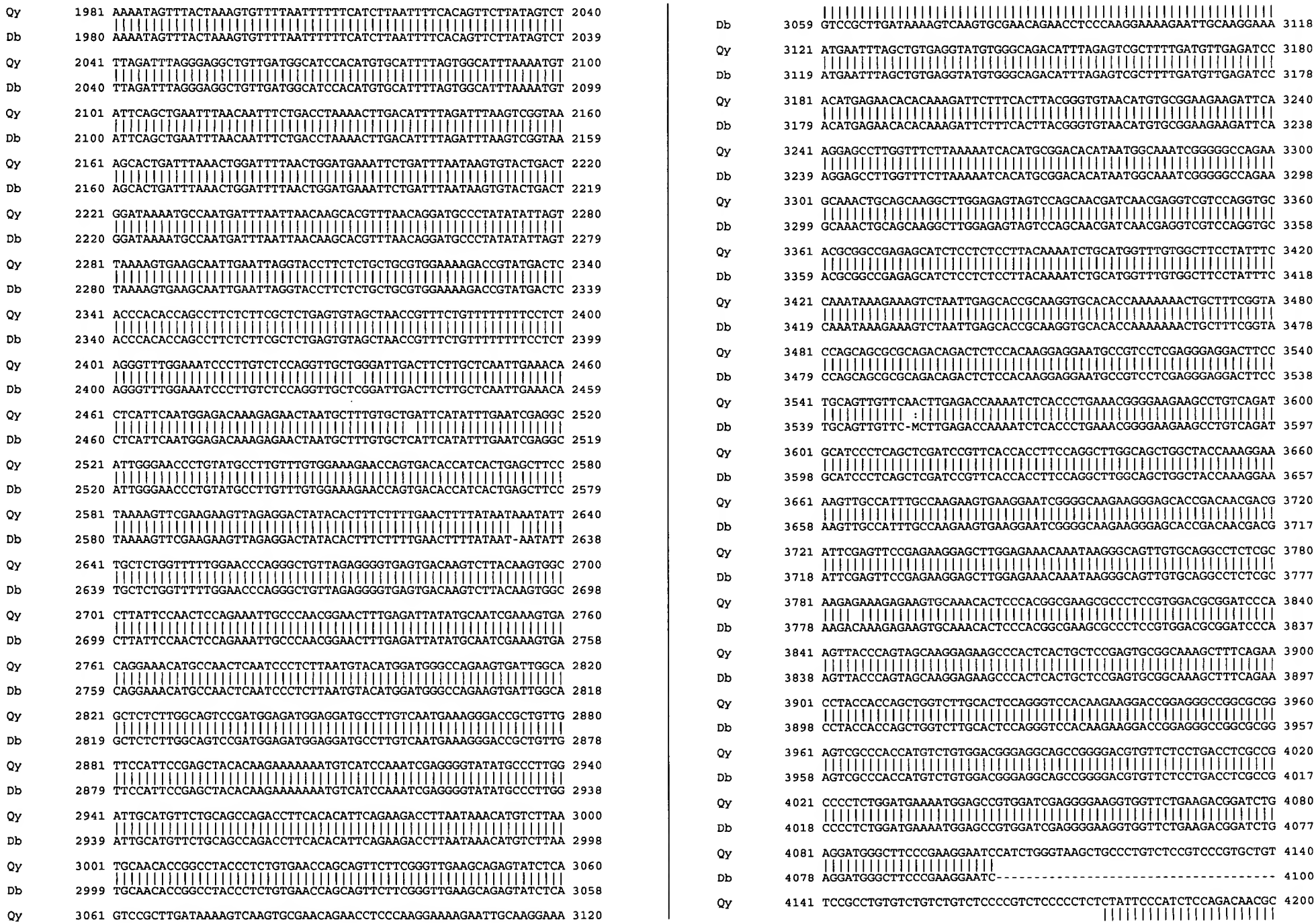

Qy 6418 CCTATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTGT 6477
Db 13948 CCTATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATTTGGGGACAGGTTTTCATTTGT 14007
Qy 6478 CAGATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTG 6537
Db 14008 CAGATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTG 14067
Qy 6538 CTAAATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAA 6597
Db 14068 CTAAATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAA 14127
Qy 6598 GATTAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCT 6657
Db 14128 GATTAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCT 14187
Qy 6658 TATTTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTA 6717
Db 14188 TATTTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTA 14247
Qy 6718 TTCTGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCT 6777
Db 14248 TTCTGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCT 14307
Qy 6778 GCCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATC 6837
Db 14308 GCCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATC 14367
Qy 6838 TTTCTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCA 6897
Db 14368 TTTCTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCA 14427
Qy 6898 GAAAAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGC 6957
Db 14428 GAAAAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGC 14487
Qy 6958 ATTTGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCAT 7017
Db 14488 ATTTGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCAT 14547
Qy 7018 GCTTTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTC 7077
Db 14548 GCTTTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTC 14607
Qy 7078 AGCTACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGA 7137
Db 14608 AGCTACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGA 14667
Qy 7138 AAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCA 7197
Db 14668 AAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCA 14727
Qy 7198 AATTATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAA 7257
Db 14728 AATTATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAA 14787
Qy 7258 CCATGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTC 7317
Db 14788 CCATGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTC 14847
Qy 7318 TGCCAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATT 7377
Db 14848 TGCCAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATT 14907
Qy 7378 TTTATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTT 7437
Db 14908 TTTATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTT 14967
Qy 7438 TCTTTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTG 7497
Db 14968 TCTTTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTG 15027


Qy 7498 GTGCAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAG 7557
Db 15028 GTGCAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAG 15087
Qy 7558 CCTCCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTT 7617
Db 15088 CCTCCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTT 15147
Qy 7618 TAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTG 7677
Db 15148 TAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTG 15207
Qy 7678 ATCCGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTG 7737
Db 15208 ATCCGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTG 15267
Qy 7738 AGCCACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTT 7797
Db 15268 AGCCACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTT 15327
Qy 7798 ATCTTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAAT 7857
Db 15328 ATCTTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAAT 15387
Qy 7858 GAAATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTA 7917
Db 15388 GAAATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTA 15447
Qy 7918 GGAGGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCA 7977
Db 15448 GGAGGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCA 15507
Qy 7978 AGTTATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATAT 8037
Db 15508 AGTTATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATAT 15567
Qy 8038 TTCCCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATT 8097
Db 15568 TTCCCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATT 15627
Qy 8098 GAAGATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGA 8157
Db 15628 GAAGATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGA 15687
Qy 8158 GAAATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAA 8217
Db 15688 GAAATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAA 15747
Qy 8218 TAAATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTT 8277
Db 15748 TAAATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTT 15807
Qy 8278 CTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAG 8337
Db 15808 CTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAG 15867
Qy 8338 ACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCT 8397
Db 15868 ACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCT 15927
Qy 8398 GAAGTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGT 8457
Db 15928 GAAGTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGT 15987
Qy 8458 GCGCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCA 8517
Db 15988 GCGCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCA 16047
Qy 8518 CCTGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTC 8577
Db 16048 CCTGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTC 16107
Qy 8578 CTCTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCT 8637

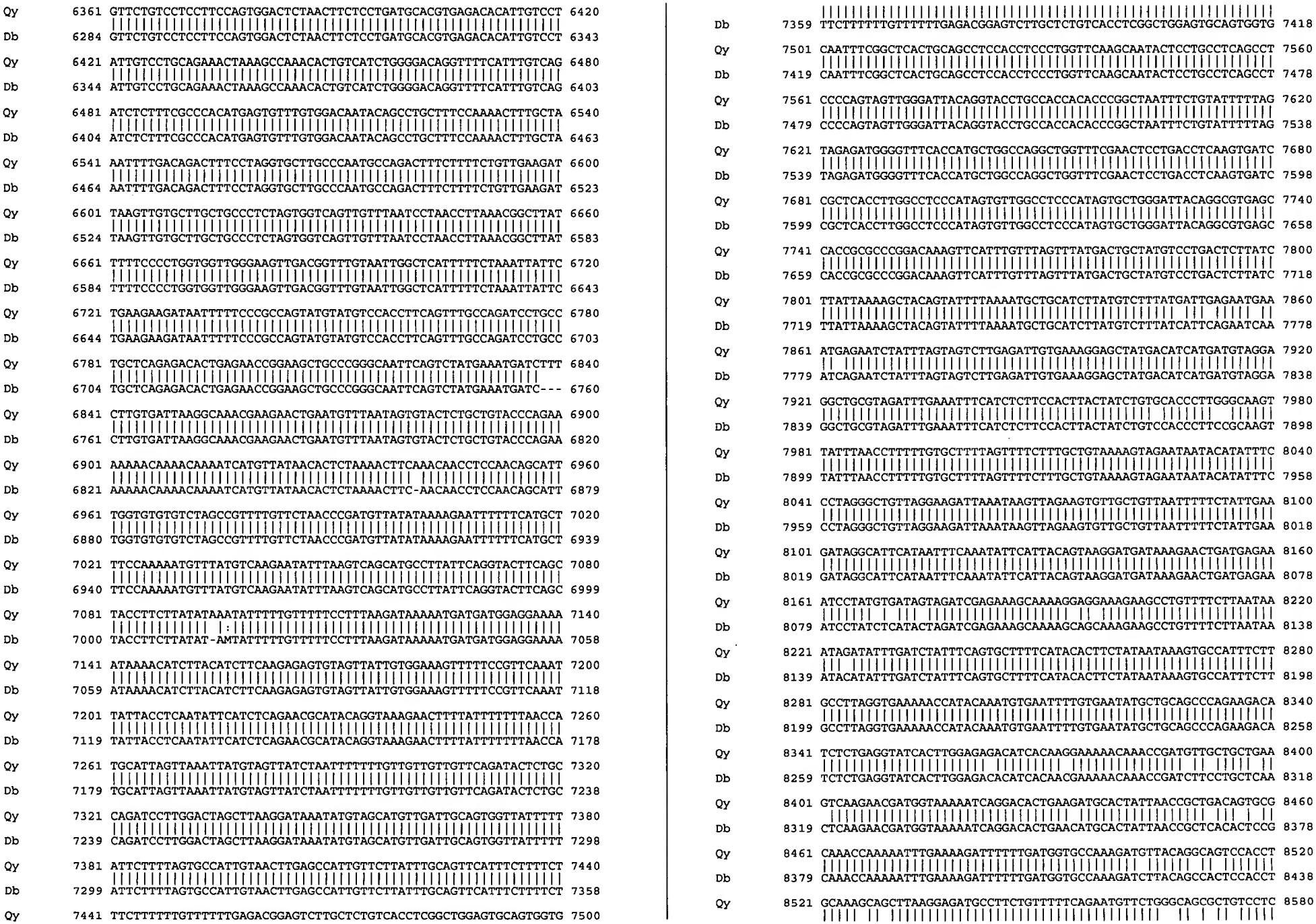
```
Db      16108 CTCTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCT 16167

Qy       8638 GATAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCA 8697

Db      16168 GATAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCA 16227

Qy       8698 GTTGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCG 8757

Db      16228 GTTGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCG 16287

Qy       8758 GATGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCA 8817

Db      16288 GATGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCA 16347

Qy       8818 GCTGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTG 8877

Db      16348 GCTGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTG 16407

Qy       8878 GGGGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACC 8937

Db      16408 GGGGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACC 16467

Qy       8938 TGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTG 8997

Db      16468 TGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTG 16527

Qy       8998 GAGCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGT 9057

Db      16528 GAGCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGT 16587

Qy       9058 CGACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTC 9117

Db      16588 CGACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCTCTAGTTTC 16647

Qy       9118 TGTAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAG 9177

Db      16648 TGTAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAG 16707

Qy       9178 GGGAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGC 9237

Db      16708 GGGAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGC 16767

Qy       9238 ACTTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGG 9297

Db      16768 ACTTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGG 16827

Qy       9298 GCCGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCG 9357

Db      16828 GCCGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCG 16887

Qy       9358 GATAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCC 9417

Db      16888 GATAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCC 16947

Qy       9418 ACCCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCG 9477

Db      16948 ACCCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCG 17007

Qy       9478 TGGGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAA 9537

Db      17008 TGGGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAA 17067

Qy       9538 TTTGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAG 9597

Db      17068 TTTGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAG 17127

Qy       9598 CCCGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATC 9657

Db      17128 CCCGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATC 17187

Qy       9658 ACATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGG 9717

Db      17188 ACATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGG 17247

Qy       9718 TTCCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGT 9777

Db      17248 TTCCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGT 17307

Qy       9778 GGCTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTA 9837

Db      17308 GGCTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTA 17367

Qy       9838 GAAGGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAAT 9897

Db      17368 GAAGGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAAT 17427

Qy       9898 CCAGGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGT 9957

Db      17428 CCAGGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGT 17487

Qy       9958 TCTTGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000

Db      17488 TCTTGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 17530
```

RESULT 2
AAV09023
ID AAV09023 standard; DNA; 10282 BP.
XX
AC AAV09023;
XX
DT 21-JUL-1998 (first entry)
XX
DE Homo sapiens 20q13 amplicon ZABC-1 genomic sequence.
XX
KW 20q13 amplicon; chromosome 20; tumour; detection; ZABC-1 gene;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentation;
KW Leber's congenital amaurosis; zinc finger amplified in breast cancer; ds.
XX
OS Homo sapiens.
XX
PN WO9802539-A1.
XX
PD 22-JAN-1998.
XX
PF 15-JUL-1997; 97WO-US012343.
XX
PR 15-JUL-1996; 96US-00680395.
PR 16-OCT-1996; 96US-00731499.
PR 17-JAN-1997; 97US-00785532.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
XX
DR WPI; 1998-110587/10.
DR P-PSDB; AAW23975.
XX
PT New sequences from the 20q13 amplicon - used for detecting chromosomal
PT abnormalities, particularly tumours, and for developing products for
PT treating diseases.
XX
PS Claim 1; Page 64-67; 91pp; English.
XX
CC The sequence is that of the genomic sequence of ZABC-1 (zinc finger
CC amplified in breast cancer). It maps to the core of the 20q13.2 amplicon
CC and is overexpressed in primary tumours and breast cancer cell lines
CC having 20q13.2 amplification. The exact coding region for the genomic
CC sequence is not given. The sequence can be used as a probe for the
CC detection of chromosomal abnormalities at 20q13. It and other sequences
CC isolated from the 20q13 amplicon are consistently amplified in primary
CC tumours. These sequences are useful as probes or as probe targets for
CC monitoring the relative copy number of corresponding sequences from a
CC biological sample such as tumour cells. The sequences can also be used in

```
CC   therapeutic applications for modulating the expression of the endogenous
CC   gene or the activity of the gene product. Examples of therapeutic
CC   approaches include antisense inhibition of gene expression, gene therapy,
CC   and monoclonal antibodies that specifically bind the gene products. The
CC   products can also be used in the treatment of other diseases, e.g. age-
CC   related macular degeneration, Leber's congenital amaurosis and retinitis
CC   pigmentation
XX
SQ   Sequence 10282 BP; 2820 A; 2222 C; 2191 G; 3047 T; 0 U; 2 Other;

  Query Match             96.5%;  Score 9651.6;  DB 2;  Length 10282;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches 9845;  Conservative    2;  Mismatches   71;  Indels   82;  Gaps    7;

Qy          1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 60
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db          1 CCATCATATTTCTTATTTTTTTCGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 60

Qy         61 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 120

Qy        121 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 180

Qy        181 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 240

Qy        241 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300

Qy        301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 360

Qy        361 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 420

Qy        421 AATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTCAAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTCAAAT 480

Qy        481 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAAAACTCCAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAACCAAAAAACTCCAC 540

Qy        541 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 600
              |||| |||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        541 TATTCCCTATTGCCACTATTTGATTTTTT-AAAAATAAGCGTATTTTAGCATCTAAAAGT 599

Qy        601 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 659

Qy        661 TTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 719

Qy        721 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 779

Qy        781 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 839

Qy        841 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 899

Qy        901 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 959

Qy        961 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 1019

Qy       1021 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 1079

Qy       1081 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 1139

Qy       1141 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 1199

Qy       1201 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 1259

Qy       1261 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 1319

Qy       1321 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 1379

Qy       1381 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 1439

Qy       1441 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 1499

Qy       1501 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAAATGAGAGGGCCCATTTTCTGATGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1500 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAAATGAGAGGGCCCATTTTCTGATGA 1559

Qy       1561 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 1619

Qy       1621 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 1679

Qy       1681 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1680 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 1739

Qy       1741 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 1799

Qy       1801 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db       1800 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTCTGCTC 1859

Qy       1861 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1860 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 1919

Qy       1921 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 1979
```


Qy 1981 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 2040
Db 1980 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 2039
Qy 2041 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 2100
Db 2040 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 2099
Qy 2101 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 2160
Db 2100 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 2159
Qy 2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 2220
Db 2160 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 2219
Qy 2221 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 2280
Db 2220 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 2279
Qy 2281 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 2340
Db 2280 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 2339
Qy 2341 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTCT 2400
Db 2340 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTCT 2399
Qy 2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 2460
Db 2400 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTCGGATTGACTTCTTGCTCAATTGAAACA 2459
Qy 2461 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520
Db 2460 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTCATTCATATTTGAATCGAGGC 2519
Qy 2521 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2580
Db 2520 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2579
Qy 2581 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 2640
Db 2580 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAAT-AATATT 2638
Qy 2641 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2700
Db 2639 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2698
Qy 2701 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2760
Db 2699 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2758
Qy 2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2820
Db 2759 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2818
Qy 2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2880
Db 2819 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2878
Qy 2881 TTCCATTCCGAGCTACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 2940
Db 2879 TTCCATTCCGAGCTACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 2938
Qy 2941 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 3000
Db 2939 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 2998
Qy 3001 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 3060
Db 2999 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 3058
Qy 3061 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 3120
Db 3059 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 3118
Qy 3121 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 3180
Db 3119 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 3178
Qy 3181 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 3240
Db 3179 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 3238
Qy 3241 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 3300
Db 3239 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 3298
Qy 3301 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 3360
Db 3299 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 3358
Qy 3361 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 3420
Db 3359 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 3418
Qy 3421 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 3480
Db 3419 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 3478
Qy 3481 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 3540
Db 3479 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 3538
Qy 3541 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 3600
Db 3539 TGCAGTTGTTC-MCTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 3597
Qy 3601 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 3660
Db 3598 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 3657
Qy 3661 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 3720
Db 3658 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 3717
Qy 3721 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 3780
Db 3718 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 3777
Qy 3781 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 3840
Db 3778 AAGACAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 3837
Qy 3841 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 3900
Db 3838 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 3897
Qy 3901 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 3960
Db 3898 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 3957
Qy 3961 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 4020
Db 3958 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 4017
Qy 4021 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 4080
Db 4018 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 4077
Qy 4081 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 4140
Db 4078 AGGATGGGCTTCCCGAAGGAATC---------- 4100
Qy 4141 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 4200

```
Db       4101 ------------------------------------CTATTCCCATCTCCAGACAACGC 4123

Qy       4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 4260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4124 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 4183

Qy       4261 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 4320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4184 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 4243

Qy       4321 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 4380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4244 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 4303

Qy       4381 TCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 4440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4304 TCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 4363

Qy       4441 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 4500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4364 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 4423

Qy       4501 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 4560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4424 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 4483

Qy       4561 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 4620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4484 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 4543

Qy       4621 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 4680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4544 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 4603

Qy       4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATCAACTAGAGCCAG 4740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4604 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATCAACTAGAGCCAG 4663

Qy       4741 GAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGCAAG 4800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4664 GAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGCAAG 4723

Qy       4801 CTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGGAAT 4860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4724 CTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGGAAT 4783

Qy       4861 GTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTGTAC 4920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4784 GTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTGTAC 4843

Qy       4921 TAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCCTGA 4980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4844 TAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCCTGA 4903

Qy       4981 GCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCATCC 5040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4904 GCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCATCC 4963

Qy       5041 TCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATCTTG 5100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4964 TCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATCTTG 5023

Qy       5101 CTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGATGG 5160
              |||||||||||||||||||||||||||||||| |||||||||||| | ||||||||
Db       5024 CTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGCCCGTAATTGACAACCCAATGATGG 5083

Qy       5161 GAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGTGCC 5220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5084 GAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGTGCC 5143

Qy       5221 TCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTATGAG 5280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5144 TCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTATGAG 5203

Qy       5281 ACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAGCCA 5340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5204 ACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAGCCA 5263

Qy       5341 GTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAAATC 5400
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5264 CTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAAATC 5323

Qy       5401 TTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAAT 5460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5324 TTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAAT 5383

Qy       5461 CTAATGTCTTAATGGCATCATAAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAAGTG 5520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5384 CTAATGTCTTAATGGCATCATAAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAAGTG 5443

Qy       5521 GTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGAAAG 5580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5444 GTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGAAAG 5503

Qy       5581 AAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGGCAG 5640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5504 AAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGGCAG 5563

Qy       5641 GACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCTACC 5700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5564 GACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCTACC 5623

Qy       5701 ACTGCGTTGTTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCCCTG 5760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5624 ACTGCGTTGTTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCCCTG 5683

Qy       5761 TTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGCCTA 5820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5684 TTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGCCTA 5743

Qy       5821 TTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCTTGA 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5744 TTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCTTGA 5803

Qy       5881 AAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGCATT 5940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5804 AAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGCATT 5863

Qy       5941 TCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGTCGG 6000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5864 TCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGTCGG 5923

Qy       6001 AAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAAAAA 6060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5924 AAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAAAAA 5983

Qy       6061 AGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAGTCT 6120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5984 AGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAGTCT 6043

Qy       6121 AGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTAAAT 6180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6044 AGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTAAAT 6103

Qy       6181 TTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTGGAT 6240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6104 TTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTGGAT 6163

Qy       6241 AGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTTTGG 6300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6164 AGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTTTGG 6223

Qy       6301 GGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTTGAA 6360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6224 GGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTTGAA 6283
```

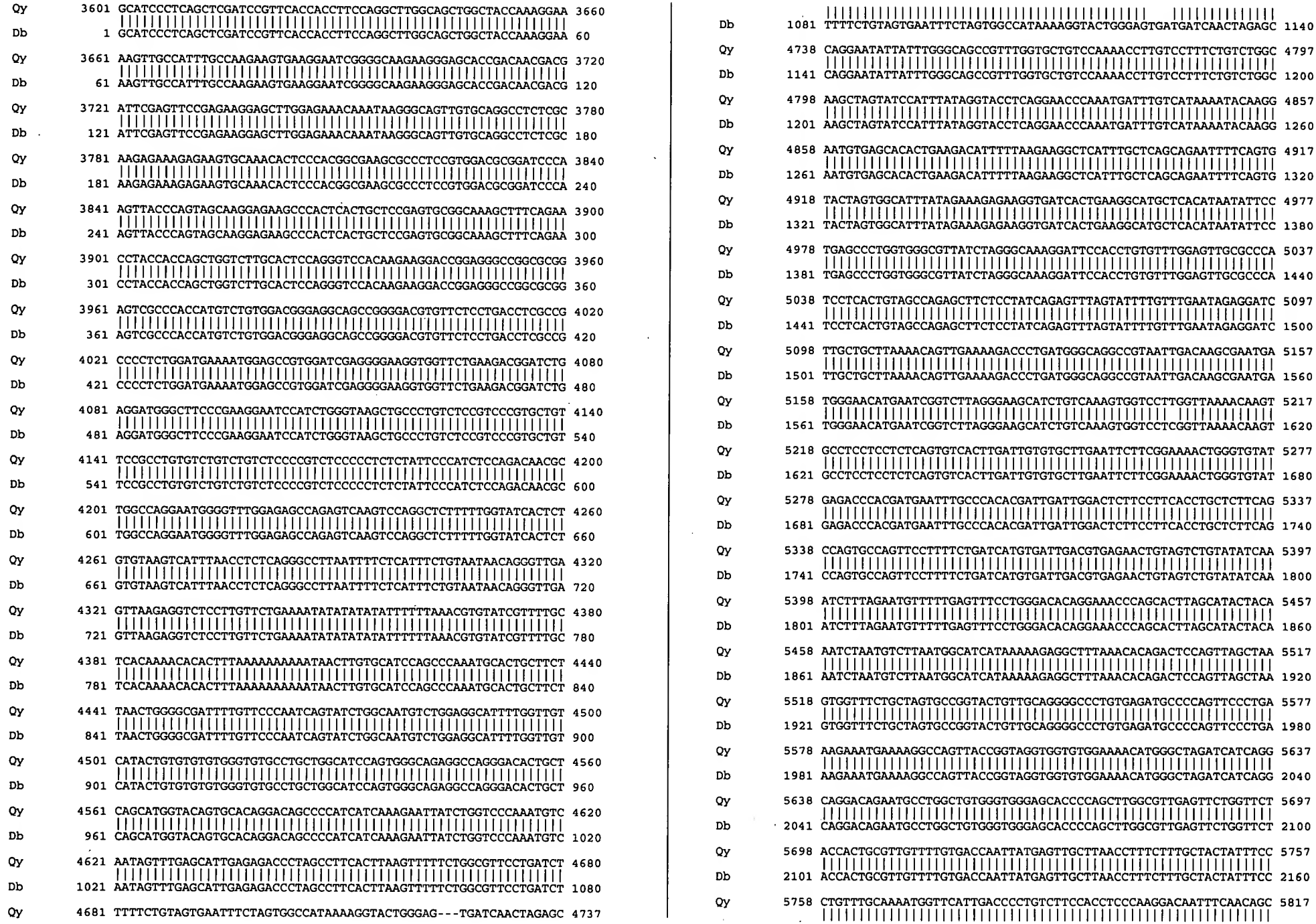

Qy 6361 GTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGTCCT 6420
Db 6284 GTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGTCCT 6343
Qy 6421 ATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTGTCAG 6480
Db 6344 ATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTGTCAG 6403
Qy 6481 ATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTGCTA 6540
Db 6404 ATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTGCTA 6463
Qy 6541 AATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAAGAT 6600
Db 6464 AATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAAGAT 6523
Qy 6601 TAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCTTAT 6660
Db 6524 TAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCTTAT 6583
Qy 6661 TTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTATTC 6720
Db 6584 TTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTATTC 6643
Qy 6721 TGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCTGCC 6780
Db 6644 TGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCTGCC 6703
Qy 6781 TGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATCTTT 6840
Db 6704 TGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATC--- 6760
Qy 6841 CTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCAGAA 6900
Db 6761 CTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCAGAA 6820
Qy 6901 AAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGCATT 6960
Db 6821 AAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTC-AACAACCTCCAACAGCATT 6879
Qy 6961 TGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCATGCT 7020
Db 6880 TGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCATGCT 6939
Qy 7021 TTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTCAGC 7080
Db 6940 TTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTCAGC 6999
Qy 7081 TACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGAAAA 7140
Db 7000 TACCTTCTTATAT-AMTATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGAAAA 7058
Qy 7141 ATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCAAAT 7200
Db 7059 ATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCAAAT 7118
Qy 7201 TATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAACCA 7260
Db 7119 TATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAACCA 7178
Qy 7261 TGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTCTGC 7320
Db 7179 TGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTCTGC 7238
Qy 7321 CAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATTTTT 7380
Db 7239 CAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATTTTT 7298
Qy 7381 ATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTTTCT 7440
Db 7299 ATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTTTCT 7358
Qy 7441 TTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTGGTG 7500
Db 7359 TTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTGGTG 7418
Qy 7501 CAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAGCCT 7560
Db 7419 CAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAGCCT 7478
Qy 7561 CCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTTTAG 7620
Db 7479 CCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTTTAG 7538
Qy 7621 TAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATC 7680
Db 7539 TAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATC 7598
Qy 7681 CGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTGAGC 7740
Db 7599 CGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTGAGC 7658
Qy 7741 CACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTTATC 7800
Db 7659 CACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTTATC 7718
Qy 7801 TTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAATGAA 7860
Db 7719 TTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATCATTCAGAATCAA 7778
Qy 7861 ATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTAGGA 7920
Db 7779 ATCAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTAGGA 7838
Qy 7921 GGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCAAGT 7980
Db 7839 GGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTCCACCCTTCCGCAAGT 7898
Qy 7981 TATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATATTTC 8040
Db 7899 TATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATATTTC 7958
Qy 8041 CCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATTGAA 8100
Db 7959 CCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATTGAA 8018
Qy 8101 GATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGAGAA 8160
Db 8019 GATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGAGAA 8078
Qy 8161 ATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAATAA 8220
Db 8079 ATCCTATCTCATACTAGATCGAGAAAGCAAAAGCAGCAAAGAAGCCTGTTTTCTTAATAA 8138
Qy 8221 ATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTTCTT 8280
Db 8139 ATACATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTTCTT 8198
Qy 8281 GCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACA 8340
Db 8199 GCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACA 8258
Qy 8341 TCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAA 8400
Db 8259 TCTCTGAGGTATCACTTGGAGACACATCACAACGAAAAACAAACCGATCTTCCTGCTCAA 8318
Qy 8401 GTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCG 8460
Db 8319 CTCAAGAACGATGGTAAAAATCAGGACACTGAACATGCACTATTAACCGCTCACACTCCG 8378
Qy 8461 CAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCT 8520
Db 8379 CAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATCTTACAGCCACTCCACCT 8438
Qy 8521 GCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTC 8580

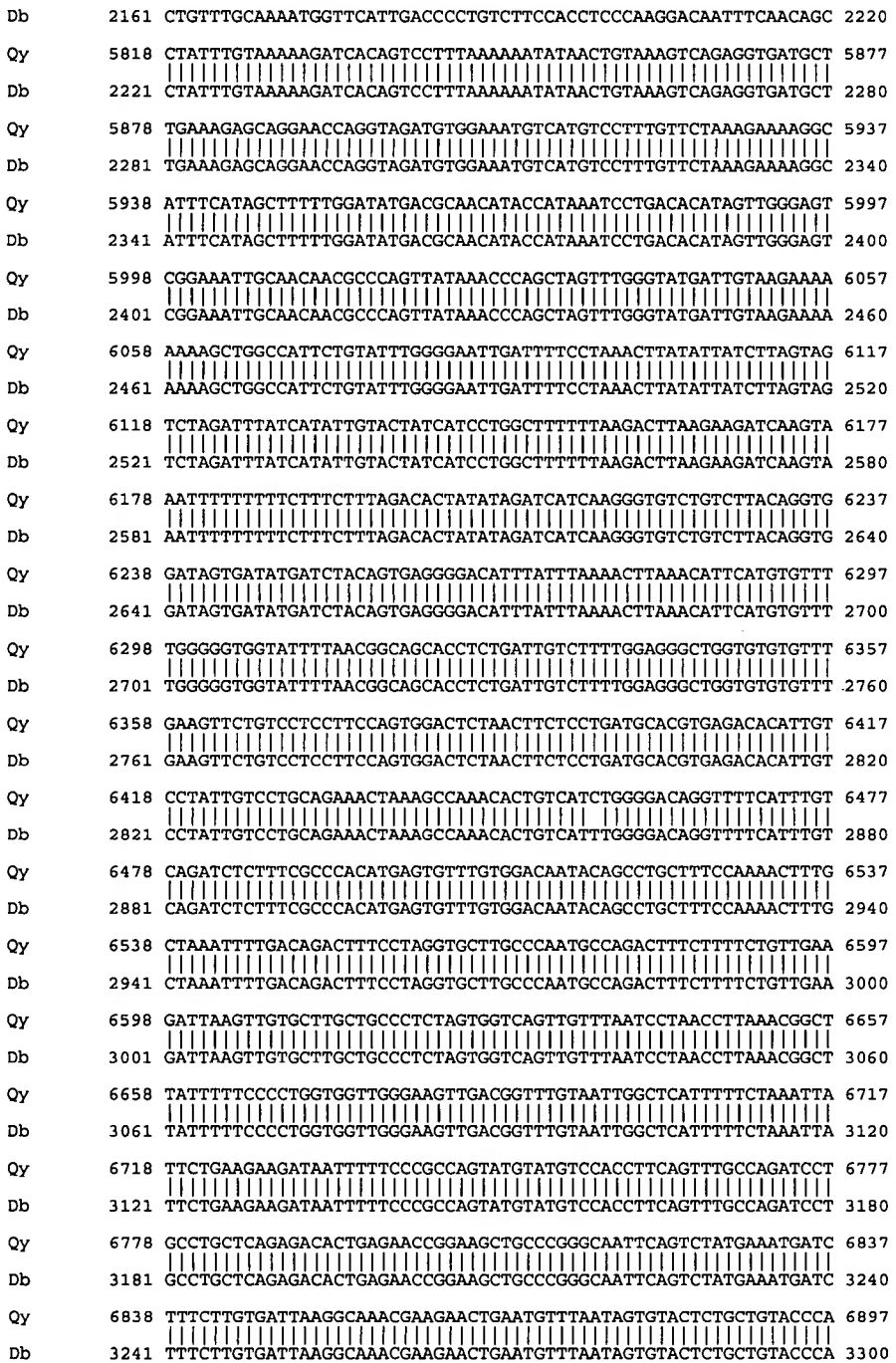
```
Db       8439 GCAAACCACCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGCCCACCCCTCTCCTC 8498

Qy       8581 TCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGAT 8640
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8499 TCACCACCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGAT 8558

Qy       8641 AAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTT 8700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8559 AAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTT 8618

Qy       8701 GAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGAT 8760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8619 GAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGAT 8678

Qy       8761 GGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCT 8820
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||||||
Db       8679 GGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGCAGACCCCAGCT 8738

Qy       8821 GACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGG 8880
             |||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db       8739 GACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAAAAAATCCGTGGGG 8798

Qy       8881 GCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGT 8940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8799 GCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGT 8858

Qy       8941 CCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAG 9000
             |||||||||||||||||||||||||||||| || ||||||| |||||||| | ||| |
Db       8859 CCATTTTGTACCTTCAAGACATTTTATCCACAACTTTTAATCATGCACCACACACTCCAC 8918

Qy       9001 CATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGA 9060
             |||||||||||||||||||||||||||||||||||||||||||||||||||| || ||||
Db       8919 CATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTACAACTCGA 8978

Qy       9061 CGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGT 9120
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8979 CGTACCCGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGT 9038

Qy       9121 AAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGG 9180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9039 AAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGG 9098

Qy       9181 AAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACT 9240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9099 AAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACT 9158

Qy       9241 TTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCC 9300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9159 TTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCC 9218

Qy       9301 GCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGAT 9360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9219 GCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGAT 9278

Qy       9361 AAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACC 9420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9279 AAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACC 9338

Qy       9421 CTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGG 9480
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db       9339 CTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACACCCCCTGC 9398

Qy       9481 GCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTT 9540
             |||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||||
Db       9399 GCACCTCCGGCAACAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTT 9458

Qy       9541 GGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCC 9600
             |||||||||||||||||||||| || ||||||||| ||||||||||||||||||||||||||
Db       9459 GGTGAGCCCCTTCCAAAAAGACTCAACTCCAGCGTGCTTGCCCTTGACGTTGACCAGCCC 9518

Qy       9601 GGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACA 9660
             |||||||||||||| || ||||||||||||||||||||||||||||||||||||||||||||
Db       9519 GGGGCCAATTACACAACAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACA 9578

Qy       9661 TCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTC 9720
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db       9579 TCACTGTTACCGCAGGACTGTGTGTATCCGTCCCAGGCGCTGCCTCCCAAACCAAGGTTC 9638

Qy       9721 CTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGC 9780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9639 CTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGC 9698

Qy       9781 TCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAA 9840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9699 TCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAA 9758

Qy       9841 GGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAATCCA 9900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9759 GGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAATCCA 9818

Qy       9901 GGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGTTCT 9960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9819 GGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGTTCT 9878

Qy       9961 TGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000
             |||||||||||||||||||||||||||||||||||||||||
Db       9879 TGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 9918
```

RESULT 3
AAK81093
ID AAK81093 standard; DNA; 14906 BP.
XX
AC AAK81093;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35905.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 35905; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 14906 BP; 3959 A; 3088 C; 3333 G; 4526 T; 0 U; 0 Other;

Query Match 63.8%; Score 6382.2; DB 4; Length 14906;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 6397; Conservative 0; Mismatches 3; Indels 3; Gaps 1;


Qy 3601 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 3660
Db 1 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 60
Qy 3661 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 3720
Db 61 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 120
Qy 3721 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 3780
Db 121 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 180
Qy 3781 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 3840
Db 181 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 240
Qy 3841 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 3900
Db 241 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 300
Qy 3901 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 3960
Db 301 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 360
Qy 3961 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 4020
Db 361 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 420
Qy 4021 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 4080
Db 421 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 480
Qy 4081 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 4140
Db 481 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 540
Qy 4141 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 4200
Db 541 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 600
Qy 4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 4260
Db 601 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 660
Qy 4261 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 4320
Db 661 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 720
Qy 4321 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATTTTTTAAACGTGTATCGTTTTGC 4380
Db 721 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATTTTTTAAACGTGTATCGTTTTGC 780
Qy 4381 TCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 4440
Db 781 TCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 840
Qy 4441 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 4500
Db 841 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 900
Qy 4501 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 4560
Db 901 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 960
Qy 4561 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 4620
Db 961 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 1020
Qy 4621 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 4680
Db 1021 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 1080
Qy 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAG---TGATCAACTAGAGC 4737
Db 1081 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATGATCAACTAGAGC 1140
Qy 4738 CAGGAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGC 4797
Db 1141 CAGGAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGC 1200
Qy 4798 AAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGG 4857
Db 1201 AAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGG 1260
Qy 4858 AATGTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTG 4917
Db 1261 AATGTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTG 1320
Qy 4918 TACTAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCC 4977
Db 1321 TACTAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCC 1380
Qy 4978 TGAGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCA 5037
Db 1381 TGAGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCA 1440
Qy 5038 TCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATC 5097
Db 1441 TCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATC 1500
Qy 5098 TTGCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGA 5157
Db 1501 TTGCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGA 1560
Qy 5158 TGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGT 5217
Db 1561 TGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTCGGTTAAAACAAGT 1620
Qy 5218 GCCTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTAT 5277
Db 1621 GCCTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTAT 1680
Qy 5278 GAGACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAG 5337
Db 1681 GAGACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAG 1740
Qy 5338 CCAGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAA 5397
Db 1741 CCAGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAA 1800
Qy 5398 ATCTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACA 5457
Db 1801 ATCTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACA 1860
Qy 5458 AATCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAA 5517
Db 1861 AATCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAA 1920
Qy 5518 GTGGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGA 5577
Db 1921 GTGGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGA 1980
Qy 5578 AAGAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGG 5637
Db 1981 AAGAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGG 2040
Qy 5638 CAGGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCT 5697
Db 2041 CAGGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCT 2100
Qy 5698 ACCACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCC 5757
Db 2101 ACCACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCC 2160
Qy 5758 CTGTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGC 5817


Db 2161 CTGTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGC 2220
Qy 5818 CTATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCT 5877
Db 2221 CTATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCT 2280
Qy 5878 TGAAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGC 5937
Db 2281 TGAAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGC 2340
Qy 5938 ATTTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGT 5997
Db 2341 ATTTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGT 2400
Qy 5998 CGGAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAA 6057
Db 2401 CGGAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAA 2460
Qy 6058 AAAAGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAG 6117
Db 2461 AAAAGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAG 2520
Qy 6118 TCTAGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTA 6177
Db 2521 TCTAGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTA 2580
Qy 6178 AATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTG 6237
Db 2581 AATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTG 2640
Qy 6238 GATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTT 6297
Db 2641 GATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTT 2700
Qy 6298 TGGGGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTT 6357
Db 2701 TGGGGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTT 2760
Qy 6358 GAAGTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGT 6417
Db 2761 GAAGTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGT 2820
Qy 6418 CCTATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTGT 6477
Db 2821 CCTATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATTTGGGGACAGGTTTTCATTTGT 2880
Qy 6478 CAGATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTG 6537
Db 2881 CAGATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTG 2940
Qy 6538 CTAAATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAA 6597
Db 2941 CTAAATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAA 3000
Qy 6598 GATTAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCT 6657
Db 3001 GATTAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCT 3060
Qy 6658 TATTTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTA 6717
Db 3061 TATTTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTA 3120
Qy 6718 TTCTGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCT 6777
Db 3121 TTCTGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCT 3180
Qy 6778 GCCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATC 6837
Db 3181 GCCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATC 3240
Qy 6838 TTTCTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCA 6897
Db 3241 TTTCTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCA 3300
Qy 6898 GAAAAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGC 6957
Db 3301 GAAAAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGC 3360
Qy 6958 ATTTGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCAT 7017
Db 3361 ATTTGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCAT 3420
Qy 7018 GCTTTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTC 7077
Db 3421 GCTTTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTC 3480
Qy 7078 AGCTACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGA 7137
Db 3481 AGCTACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGA 3540
Qy 7138 AAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCA 7197
Db 3541 AAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCA 3600
Qy 7198 AATTATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAA 7257
Db 3601 AATTATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAA 3660
Qy 7258 CCATGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTC 7317
Db 3661 CCATGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTC 3720
Qy 7318 TGCCAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATT 7377
Db 3721 TGCCAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATT 3780
Qy 7378 TTTATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTT 7437
Db 3781 TTTATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTT 3840
Qy 7438 TCTTTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTG 7497
Db 3841 TCTTTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTG 3900
Qy 7498 GTGCAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAG 7557
Db 3901 GTGCAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAG 3960
Qy 7558 CCTCCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTT 7617
Db 3961 CCTCCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTT 4020
Qy 7618 TAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTG 7677
Db 4021 TAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTG 4080
Qy 7678 ATCCGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTG 7737
Db 4081 ATCCGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTG 4140
Qy 7738 AGCCACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTT 7797
Db 4141 AGCCACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTT 4200
Qy 7798 ATCTTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAAT 7857
Db 4201 ATCTTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAAT 4260
Qy 7858 GAAATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTA 7917
Db 4261 GAAATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTA 4320
Qy 7918 GGAGGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCA 7977
Db 4321 GGAGGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCA 4380

```
Qy       7978 AGTTATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATAT 8037
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4381 AGTTATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATAT 4440

Qy       8038 TTCCCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATT 8097
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4441 TTCCCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATT 4500

Qy       8098 GAAGATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGA 8157
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4501 GAAGATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGA 4560

Qy       8158 GAAATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAA 8217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4561 GAAATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAA 4620

Qy       8218 TAAATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTT 8277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 TAAATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTT 4680

Qy       8278 CTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAG 8337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4681 CTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAG 4740

Qy       8338 ACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCT 8397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 ACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCT 4800

Qy       8398 GAAGTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGT 8457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4801 GAAGTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGT 4860

Qy       8458 GCGCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCA 8517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4861 GCGCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCA 4920

Qy       8518 CCTGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTC 8577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4921 CCTGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTC 4980

Qy       8578 CTCTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAAATGCAGCTGATGACAGTGCT 8637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 CTCTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAAATGCAGCTGATGACAGTGCT 5040

Qy       8638 GATAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCA 8697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5041 GATAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCA 5100

Qy       8698 GTTGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCG 8757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5101 GTTGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCG 5160

Qy       8758 GATGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCA 8817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5161 GATGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCA 5220

Qy       8818 GCTGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTG 8877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5221 GCTGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTG 5280

Qy       8878 GGGGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACC 8937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5281 GGGGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACC 5340

Qy       8938 TGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTG 8997
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5341 TGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTG 5400

Qy       8998 GAGCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGT 9057
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5401 GAGCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGT 5460

Qy       9058 CGACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTC 9117
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       5461 CGACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCTCTAGTTTC 5520

Qy       9118 TGTAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAG 9177
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5521 TGTAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAG 5580

Qy       9178 GGGAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGC 9237
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5581 GGGAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGC 5640

Qy       9238 ACTTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGG 9297
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5641 ACTTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGG 5700

Qy       9298 GCCGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCG 9357
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5701 GCCGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCG 5760

Qy       9358 GATAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCC 9417
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5761 GATAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCC 5820

Qy       9418 ACCCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCG 9477
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5821 ACCCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCG 5880

Qy       9478 TGGGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAA 9537
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5881 TGGGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAA 5940

Qy       9538 TTTGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAG 9597
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5941 TTTGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAG 6000

Qy       9598 CCCGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATC 9657
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6001 CCCGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATC 6060

Qy       9658 ACATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGG 9717
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6061 ACATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGG 6120

Qy       9718 TTCCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGT 9777
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6121 TTCCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGT 6180

Qy       9778 GGCTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTA 9837
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6181 GGCTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTA 6240

Qy       9838 GAAGGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAAT 9897
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6241 GAAGGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAAT 6300

Qy       9898 CCAGGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGT 9957
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6301 CCAGGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGT 6360

Qy       9958 TCTTGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000
             |||||||||||||||||||||||||||||||||||||||||||
Db       6361 TCTTGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 6403
```

RESULT 4
ACC50334
ID ACC50334 standard; cDNA; 5632 BP.
XX
AC ACC50334;
XX
DT 12-JUN-2003 (first entry)
XX
DE Breast cancer associated cDNA sequence SEQ ID NO:505.

```
XX
KW   Human; breast cancer; cytostatic; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003004989-A2.
XX
PD   16-JAN-2003.
XX
PF   21-JUN-2002; 2002WO-US019669.
XX
PR   21-JUN-2001; 2001US-0299887P.
PR   27-JUN-2001; 2001US-0301572P.
PR   18-JUL-2001; 2001US-0306501P.
PR   25-SEP-2001; 2001US-0325002P.
PR   05-MAR-2002; 2002US-0362585P.
PR   14-MAY-2002; 2002US-0380391P.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Lillie J,  Gannavarapu M,  Glatt K,  Hoersh S,  Kamatkar S;
PI   Mertens M,  Monahan JE,  Myer V,  Wang Y,  Xu Y,  Zhao X,  Meyers RE;
PI   Bast RC,  Hortobagyi GN,  Pusztai L,  Meric F,  Sahin A,  Mills GB;
XX
DR   WPI; 2003-210381/20.
DR   P-PSDB; ABR47632.
XX
PT   Breast cancer diagnosis or treatment by comparing the level of expression
PT   of a marker in a patient sample with that in the control non-breast
PT   cancer sample.
XX
PS   Claim 1; SEQ ID NO 505; 128pp; English.
XX
CC   The present invention describes a method for assessing whether a patient
CC   is afflicted with breast cancer. The method comprises comparing the level
CC   of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC   ABR47386 to ABR47632) in a patient sample and the normal level of
CC   expression of the marker in a control non-breast cancer sample, where a
CC   significant increase in the level of expression of the marker in the
CC   patient sample and the normal level is an indication that the patient is
CC   afflicted with breast cancer. The breast cancer associated sequences from
CC   the present invention have cytostatic activities and can be used in gene
CC   therapy. The method is useful for diagnosing and treating breast cancer.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;

  Query Match             16.4%;  Score 1639.4;  DB 8;  Length 5632;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1640;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 2533
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 60

Qy       2534 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 2593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 120

Qy       2594 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 2653
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 180

Qy       2654 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 2713
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 240

Qy       2714 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300

Qy       2774 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 2833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 360

Qy       2834 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 2893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 420

Qy       2894 TACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 2953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 TACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 480

Qy       2954 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 540

Qy       3014 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600

Qy       3074 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 660

Qy       3134 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 3193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 720

Qy       3194 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 3253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 780

Qy       3254 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 3313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 840

Qy       3314 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 900

Qy       3374 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 960

Qy       3434 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 1020

Qy       3494 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 1080

Qy       3554 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 3613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 1140

Qy       3614 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 3673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 1200

Qy       3674 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 3733
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 1260

Qy       3734 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 3793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 1320

Qy       3794 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 3853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 1380

Qy       3854 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 3913
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 1440

Qy       3914 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 3973
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 1500

Qy       3974 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCCTCTGGATGA 4033
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCCTCTGGATGA 1560

Qy       4034 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 4093
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 1620

Qy       4094 CGAAGGAATCCATCTGGGTAA 4114
                |||||||||||||||| |||
Db       1621 CGAAGGAATCCATCTGGATAA 1641
```

RESULT 5
ADD14635
ID ADD14635 standard; cDNA; 5632 BP.
XX
AC ADD14635;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human src biomarker polynucleotide SEQ ID NO:29.
XX
KW predictor set; protein tyrosine kinase activity modulator;
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW gene therapy; drug sensitivity; genetic profile; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003062395-A2.
XX
PD 31-JUL-2003.
XX
PF 17-JAN-2003; 2003WO-US001981.
XX
PR 18-JAN-2002; 2002US-0350061P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Huang F, Fairchild CR, Lee FY, Shaw P;
XX
DR WPI; 2003-636735/60.
DR P-PSDB; ADD14040.
XX
PT New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.
XX
PS Claim 2; SEQ ID NO 29; 139pp; English.
XX
CC The present invention describes a predictor set comprising a plurality of
CC polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of
CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized
CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;

  Query Match   16.4%; Score 1639.4; DB 10; Length 5632;
  Best Local Similarity   99.9%; Pred. No. 0;
  Matches 1640; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

```
Qy       2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 2533
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 60

Qy       2534 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 2593
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 120

Qy       2594 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 2653
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 180

Qy       2654 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 2713
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 240

Qy       2714 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300

Qy       2774 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 2833
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 360

Qy       2834 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 2893
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 420

Qy       2894 TACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 2953
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 480

Qy       2954 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 3013
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 540

Qy       3014 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600

Qy       3074 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 3133
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 660

Qy       3134 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 3193
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 720

Qy       3194 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 3253
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 780

Qy       3254 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 3313
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 840
```

```
Qy       3314 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 900

Qy       3374 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 960

Qy       3434 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 1020

Qy       3494 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 1080

Qy       3554 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 3613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 1140

Qy       3614 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 3673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 1200

Qy       3674 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 3733
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 1260

Qy       3734 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 3793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 1320

Qy       3794 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 3853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 1380

Qy       3854 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 3913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 1440

Qy       3914 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 3973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 1500

Qy       3974 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 4033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 1560

Qy       4034 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 4093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 1620

Qy       4094 CGAAGGAATCCATCTGGGTAA 4114
             |||||||||||||||| |||
Db        1621 CGAAGGAATCCATCTGGATAA 1641
```

RESULT 6
ADR66761
ID AC DT DE KW KW OS XX

ADR66761 standard; DNA; 5632 BP.

ADR66761;

02-DEC-2004 (first entry)

Human prostatic carcinoma derived DNA SEQ ID 54 #4.

human; cytostatic; diagnosis; prostatic cancer;
differential expression analysis; ds.

Homo sapiens.

PN WO2004076614-A2.
PD 10-SEP-2004.
PF 22-FEB-2004; 2004WO-DE000433.
PR 27-FEB-2003; 2003DE-01009985.
PR 14-MAY-2003; 2003DE-01022134.
PA (HINZ/) HINZMANN B.
PA (DAHL/) DAHL E.
PA (ROSE/) ROSENTHAL A.
PA (HERM/) HERMANN K.
PA (PILA/) PILARSKY C.
PI Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;
PI Schmitt A, Beckmann G, Bruemmendorf T, Kinnemann H, Roepcke S;
PI Xinzhong L, Staub E;
DR WPI; 2004-653386/63.
PT New nucleic acids, and encoded proteins, from prostatic cancer tissue,
PT useful for diagnosis, treatment and in screening for specific binding
PT agents.
PS Claim 1; Page 1423-1424; 1607pp; German.

CC This invention describes novel cytostatic polynucleotide and polypeptide
CC sequences which can be used in a method for diagnosing prostatic cancer
CC or the risk of developing prostatic cancer. Diagnosis is based on
CC determining over transcription or over expression of the sequences in
CC prostatic tissue. Screening for inhibitors of the sequences or detection
CC substances involves a binding assay, any compounds that bind are
CC selected, optionally after deconvolution of mixtures. Detection of a
CC predetermined minimum level of the reporter indicates the presence of
CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
CC short-interfering RNA or ribozymes; an organic molecule of molecular
CC weight below 5000, preferably 300, that binds to the polypeptide; an
CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
CC polypeptide, preferably humanised or human; an anti-idiotype, non-human
CC (monoclonal) antibody directed against Ab or any of the above derivatised
CC with a reporter group, cell toxin, immunostimulatory molecules and/or
CC radioisotope. The polynucleotides are identified in human prostatic
CC cancer by differential expression analysis, using DNA microarrays,
CC between normal and tumorous tissues, with (over)expression being detected
CC by quantitative PCR. Analysis of prostatic cancer samples showed that
CC CD24 was upregulated in many of them. Sections of tissue, isolated from
CC prostatic cancer patients, or subjects at risk, were incubated
CC sequentially with anti-human CD4 murine monoclonal antibodies;
CC biotinylated second antibody; streptavidin-conjugated horseradish
CC peroxidase and then diaminobenzidine as colour former (brown). The
CC samples were counterstained with hemalum (blue). Malignant cells stained
CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
CC lymph node metastases were also stained. ADR65805-ADR66954 represent the
CC polynucleotide and polypeptide sequences used in the method of the
CC invention.

SQ Sequence 5632 BP; 1637 A; 1167 C; 1257 G; 1571 T; 0 U; 0 Other;

Query Match 16.4%; Score 1639.4; DB 13; Length 5632;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 2533
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 60

Qy       2534 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 2593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 120

Qy       2594 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 2653
```

```
Db      121 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 180
Qy     2654 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 2713
Db      181 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 240
Qy     2714 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
Db      241 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300
Qy     2774 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 2833
Db      301 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 360
Qy     2834 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 2893
Db      361 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 420
Qy     2894 TACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 2953
Db      421 TACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 480
Qy     2954 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 3013
Db      481 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 540
Qy     3014 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
Db      541 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600
Qy     3074 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 3133
Db      601 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 660
Qy     3134 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 3193
Db      661 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 720
Qy     3194 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 3253
Db      721 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 780
Qy     3254 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 3313
Db      781 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 840
Qy     3314 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 3373
Db      841 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 900
Qy     3374 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 3433
Db      901 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 960
Qy     3434 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 3493
Db      961 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 1020
Qy     3494 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 3553
Db     1021 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 1080
Qy     3554 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 3613
Db     1081 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 1140
Qy     3614 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 3673
Db     1141 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 1200
Qy     3674 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 3733
Db     1201 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 1260
Qy     3734 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 3793
Db     1261 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 1320
Qy     3794 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 3853
Db     1321 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 1380
Qy     3854 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 3913
Db     1381 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 1440
Qy     3914 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 3973
Db     1441 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 1500
Qy     3974 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 4033
Db     1501 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 1560
Qy     4034 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 4093
Db     1561 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 1620
Qy     4094 CGAAGGAATCCATCTGGGTAA 4114
Db     1621 CGAAGGAATCCATCTGGATAA 1641
```

RESULT 7
ADR65858
ID ADR65858 standard; DNA; 5632 BP.
XX
AC ADR65858;
XX
DT 02-DEC-2004 (first entry)
XX
DE Human prostatic carcinoma derived DNA SEQ ID 54 #1.
XX
KW human; cytostatic; diagnosis; prostatic cancer;
KW differential expression analysis; ds.
XX
OS Homo sapiens.
XX
PN WO2004076614-A2.
XX
PD 10-SEP-2004.
XX
PF 22-FEB-2004; 2004WO-DE000433.
XX
PR 27-FEB-2003; 2003DE-01009985.
PR 14-MAY-2003; 2003DE-01022134.
XX
PA (HINZ/) HINZMANN B.
PA (DAHL/) DAHL E.
PA (ROSE/) ROSENTHAL A.
PA (HERM/) HERMANN K.
PA (PILA/) PILARSKY C.
XX
PI Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;
PI Schmitt A, Beckmann G, Bruemmendorf T, Kinnemann H, Roepcke S;
PI Xinzhong L, Staub E;
XX
DR WPI; 2004-653386/63.
XX
PT New nucleic acids, and encoded proteins, from prostatic cancer tissue,
PT useful for diagnosis, treatment and in screening for specific binding
PT agents.
XX
PS Claim 1; Page 253-254; 1607pp; German.
XX

CC This invention describes novel cytostatic polynucleotide and polypeptide
CC sequences which can be used in a method for diagnosing prostatic cancer
CC or the risk of developing prostatic cancer. Diagnosis is based on
CC determining over transcription or over expression of the sequences in
CC prostatic tissue. Screening for inhibitors of the sequences or detection
CC substances involves a binding assay, any compounds that bind are
CC selected, optionally after deconvolution of mixtures. Detection of a
CC predetermined minimum level of the reporter indicates the presence of
CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
CC short-interfering RNA or ribozymes; an organic molecule of molecular
CC weight below 5000, preferably 300, that binds to the polypeptide; an
CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
CC polypeptide, preferably humanised or human; an anti-idiotype, non-human
CC (monoclonal) antibody directed against Ab or any of the above derivatised
CC with a reporter group, cell toxin, immunostimulatory molecules and/or
CC radioisotope. The polynucleotides are identified in human prostatic
CC cancer by differential expression analysis, using DNA microarrays,
CC between normal and tumorous tissues, with (over)expression being detected
CC by quantitative PCR. Analysis of prostatic cancer samples showed that
CC CD24 was upregulated in many of them. Sections of tissue, isolated from
CC prostatic cancer patients, or subjects at risk, were incubated
CC sequentially with anti-human CD4 murine monoclonal antibodies;
CC biotinylated second antibody; streptavidin-conjugated horseradish
CC peroxidase and then diaminobenzidine as colour former (brown). The
CC samples were counterstained with hemalum (blue). Malignant cells stained
CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
CC lymph node metastases were also stained. ADR65805-ADR66954 represent the
CC polynucleotide and polypeptide sequences used in the method of the
CC invention.
XX
SQ Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;

Query Match 16.4%; Score 1639.4; DB 13; Length 5632;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 2533
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 60

Qy       2534 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 2593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 120

Qy       2594 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 2653
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 180

Qy       2654 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 2713
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 240

Qy       2714 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300

Qy       2774 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 2833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 360

Qy       2834 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 2893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 420

Qy       2894 TACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 2953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 480

Qy       2954 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 540

Qy       3014 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600

Qy       3074 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 660

Qy       3134 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 3193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 720

Qy       3194 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 3253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 780

Qy       3254 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 3313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 840

Qy       3314 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 900

Qy       3374 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 960

Qy       3434 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 1020

Qy       3494 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 1080

Qy       3554 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 3613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 1140

Qy       3614 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 3673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 1200

Qy       3674 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 3733
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 1260

Qy       3734 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 3793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 1320

Qy       3794 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 3853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 1380

Qy       3854 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 3913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 1440

Qy       3914 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 3973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 1500

Qy       3974 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 4033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 1560

Qy       4034 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 4093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 1620

Qy       4094 CGAAGGAATCCATCTGGGTAA 4114
```

```
                    |||||||||||||||||| |||
Db       1621 CGAAGGAATCCATCTGGATAA 1641

RESULT 8
ADP07267
ID   ADP07267 standard; DNA; 5653 BP.
XX
AC   ADP07267;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human ZNF217 DNA.
XX
KW   ds; proliferative disease; breast; methylation; CpG; bisulfite; human.
XX
OS   Homo sapiens.
XX
PN   DE10255104-A1.
XX
PD   11-MAR-2004.
XX
PF   26-NOV-2002; 2002DE-01055104.
XX
PR   27-AUG-2002; 2002DE-01039313.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Maier S;
XX
DR   WPI; 2004-284340/27.
DR   GENBANK; NM_006526.
XX
PT   Analyzing proliferative diseases of breast cells, useful e.g. for
PT   diagnosis, prognosis and treatment of breast cancer, by determining
PT   methylation status of specific genes.
XX
PS   Claim 1; Page; 22pp; German.
XX
CC   This invention describes a novel method of analysing proliferative
CC   diseases of breast cells by determining the methylation status of certain
CC   genes. The invention also describes nucleic acids or their complements,
CC   oligomers, especially oligonucleotides or peptide nucleic acid oligomers,
CC   that hybridise to, or are identical with, any of the nucleic acids, the
CC   preparation of an oligomer array for analysing proliferative diseases of
CC   breast cells that are associated with the methylation status of CpG
CC   dinucleotides of the genes by bonding at least one oligomer to a solid
CC   phase and a kit comprising a bisulfite reagent and the oligomer. The
CC   nucleic acids are genomic sequences (5'- and/or regulatory and/or CpG-
CC   rich regions). The base sequence of the oligomer includes at least one
CC   CpG island, especially with C in the middle third of the sequence. The
CC   process involves a genomic DNA sample treated chemically, specifically
CC   with a bisulfite reagent, to convert non-methylated C to uracil or some
CC   other base having base-pairing properties different from those of C.
CC   Fragments of the treated DNA are then amplified, using primers and a
CC   polymerase and the methylation status of the genomic CpG dinucleotides is
CC   determined by analysis of the amplicons, particularly by hybridisation to
CC   the oligomer, optionally with extension of the hybridised oligomer by at
CC   least one base, or detection is by sequencing. The amplification may use
CC   methylation-specific primers. Alternatively, genomic DNA is extracted
CC   from a sample and digested with methylation-specific restriction enzymes,
CC   then the digestion fragments detected, optionally after amplification. In
CC   either method, more than 10 fragments of 100-200 bp are amplified in a
CC   single reaction vessel, using a heat-resistant DNA polymerase in PCR. The
CC   amplicons carry detectable markers, e.g. fluorophores, radioisotopes
CC   and/or releasable fragments of known mass that can be detected by mass
CC   spectrometry. The method is used for characterisation, classification,
CC   differentiation, staging, phase-estimation, diagnosis and/or therapy of
CC   proliferative diseases of breast cells. The method provides very specific
CC   classification of proliferative diseases, allowing better treatment. It
CC   can both characterise methylation status and detect single-nucleotide
CC   polymorphisms. This sequence represents human gene used to illustrate the
CC   method of the invention. NOTE: This sequence does not appear in the
CC   printed specification but has been retrieved from Genbank.
XX
SQ   Sequence 5653 BP; 1656 A; 1167 C; 1258 G; 1572 T; 0 U; 0 Other;

  Query Match             16.4%;  Score 1639.4;  DB 12;  Length 5653;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1640;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 2533
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 60

Qy       2534 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 2593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 120

Qy       2594 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 2653
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 180

Qy       2654 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 2713
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 240

Qy       2714 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300

Qy       2774 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 2833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 360

Qy       2834 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 2893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 420

Qy       2894 TACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 2953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 480

Qy       2954 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 540

Qy       3014 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600

Qy       3074 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 660

Qy       3134 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 3193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 720

Qy       3194 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 3253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 780

Qy       3254 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 3313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 840

Qy       3314 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 900

Qy       3374 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 960

Qy       3434 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 3493
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 1020

Qy       3494 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 3553
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 1080

Qy       3554 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 3613
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 1140

Qy       3614 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 3673
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 1200

Qy       3674 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 3733
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 1260

Qy       3734 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 3793
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 1320

Qy       3794 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 3853
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 1380

Qy       3854 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 3913
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 1440

Qy       3914 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 3973
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 1500

Qy       3974 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 4033
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 1560

Qy       4034 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 4093
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 1620

Qy       4094 CGAAGGAATCCATCTGGGTAA 4114
                |||||||||||||||| |||
Db       1621 CGAAGGAATCCATCTGGATAA 1641


RESULT 9
AAK61822
ID   AAK61822 standard; cDNA; 1797 BP.
XX
AC   AAK61822;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:6882.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
```

PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
DR P-PSDB; AAM89041.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Claim 1; SEQ ID NO 6882; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 1797 BP; 480 A; 377 C; 429 G; 499 T; 0 U; 12 Other;

Query Match 16.3%; Score 1633.4; DB 4; Length 1797;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1661; Conservative 4; Mismatches 13; Indels 4; Gaps 2;

```
Qy      4384 CAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCTTAA 4443
             | |  ||| | |   ||: ||||||||||||||||||||||||||||||||||||||||
Db        27 CGACCCACGCGTCCGAAWNAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCTTAA 86

Qy      4444 CTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGTCAT 4503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        87 CTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGTCAT 146

Qy      4504 ACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCTCAG 4563
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       147 ACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGTTCAG 206

Qy      4564 CATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTCAAT 4623
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       207 CATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTCAAT 266

Qy      4624 AGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCTTTT 4683
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       267 AGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCTTTT 326

Qy      4684 TCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAG---TGATCAACTAGAGCCAG 4740
             ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db       327 TCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATGATCAACTAGAGCCAG 386

Qy      4741 GAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGCAAG 4800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       387 GAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGCAAG 446

Qy      4801 CTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGGAAT 4860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       447 CTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGGAAT 506

Qy      4861 GTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTGTAC 4920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       507 GTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTGTAC 566

Qy      4921 T-AGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCCTG 4979
             | |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       567 TAAGTGGCATTTATAGAAAGAGAANGTGATCACTGAAGGCATGCTCACATAATATTCCTG 626

Qy      4980 AGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCATC 5039
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       627 AGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCATC 686

Qy      5040 CTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATCTT 5099
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       687 CTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATCTT 746

Qy      5100 GCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGATG 5159
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        747 GCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGATG 806

Qy       5160 GGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGTGC 5219
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        807 GGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTCGGTTAAAACAAGTGC 866

Qy       5220 CTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTATGA 5279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 CTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTATGA 926

Qy       5280 GACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAGCC 5339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 GACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAGCC 986

Qy       5340 AGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAAAT 5399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAAAT 1046

Qy       5400 CTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAA 5459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAA 1106

Qy       5460 TCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAAGT 5519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 TCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAAGT 1166

Qy       5520 GGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGAAA 5579
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1167 GGTTTCTGCTAGTGCCGGTACTGNTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGAAA 1226

Qy       5580 GAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGGCA 5639
              |||||||||||||:|||||||||||||||||||||:||||||||||||||||||||||||
Db       1227 GAAATGAAAAGGCCAGTTACCGGTAGGTGGTGKGGAAAACATGGGCTAGATCATCAGGCA 1286

Qy       5640 GGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCTAC 5699
              |||||||||||||:|:|||||||||||||||||||||||||||||||||||||||||||
Db       1287 GGACAGAATGCCTGGSTKTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCTAC 1346

Qy       5700 CACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCCCT 5759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 CACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCCCT 1406

Qy       5760 GTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGCCT 5819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 GTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGCCT 1466

Qy       5820 ATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCTTG 5879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 ATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCTTG 1526

Qy       5880 AAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGCAT 5939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1527 AAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGCAT 1586

Qy       5940 TTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGTCG 5999
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1587 TTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGTCG 1646

Qy       6000 GAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAAAA 6059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1647 GAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAAAA 1706

Qy       6060 AA 6061
              ||
Db       1707 AA 1708


RESULT 10
ACN44987
ID   ACN44987 standard; cDNA; 5252 BP.
XX
AC   ACN44987;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human mRNA sequence hCT1950762.
XX
KW   Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003073826-A2.
XX
PD   12-SEP-2003.
XX
PF   28-FEB-2003; 2003WO-US006235.
XX
PR   01-MAR-2002; 2002US-00087192.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
XX
DR   WPI; 2003-328604/31.
XX
PT   Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT   comprises a nucleotide sequence.
XX
PS   Claim 1; SEQ ID NO 1709; 0pp; English.
XX
CC   The present invention relates to novel DNA and protein sequences which
CC   are associated with carcinomas. The sequences are useful for: (i) for
CC   screening drug candidates; (ii) for screening of bioactive agent capable
CC   of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC   a bioactive agent capable of modulating the activity of CAP; (iv) for
CC   evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC   carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC   carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC   (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC   determining Carcinoma Associated (CA) gene copy number. In addition, the
CC   CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC   carcinoma including lymphoma. The present sequence is one such CA gene
CC   sequence. Note: This patent is an equivalent to basic patent
CC   US2002182586A1, for which no sequence data was published
XX
SQ   Sequence 5252 BP; 1507 A; 1101 C; 1177 G; 1467 T; 0 U; 0 Other;

  Query Match             15.6%;  Score 1559.6;  DB 11;  Length 5252;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1565;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       8286 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACATCTCT 8345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1506 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACATCTCT 1565

Qy       8346 GAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 8405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1566 GAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 1625

Qy       8406 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCGCAAAC 8465
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1626 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCGCAAAC 1685

Qy       8466 CAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCTGCAAA 8525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1686 CAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCTGCAAA 1745

Qy       8526 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTCTCACC 8585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1746 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTCTCACC 1805

Qy       8586 AGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGATAAAGT 8645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1806 AGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGATAAAGT 1865

Qy       8646 GAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAC 8705
```

```
Db       1866 GAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAC 1925

Qy       8706 TCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGATGGCAG 8765

Db       1926 TCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGATGGCAG 1985

Qy       8766 TACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCTGACTG 8825

Db       1986 TACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCTGACTG 2045

Qy       8826 CAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCT 8885

Db       2046 CAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCT 2105

Qy       8886 TCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGTCCATT 8945

Db       2106 TCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGTCCATT 2165

Qy       8946 TTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAGCATAA 9005

Db       2166 TTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAGCATAA 2225

Qy       9006 ATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGACGTAC 9065

Db       2226 ATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGACGTAC 2285

Qy       9066 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGTAAACC 9125

Db       2286 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCTCTAGTTTCTGTAAACC 2345

Qy       9126 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 9185

Db       2346 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 2405

Qy       9186 GAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 9245

Db       2406 GAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 2465

Qy       9246 CCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCCGCCAC 9305

Db       2466 CCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCCGCCAC 2525

Qy       9306 CAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGATAAGAC 9365

Db       2526 CAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGATAAGAC 2585

Qy       9366 AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACCCTCGG 9425

Db       2586 AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACCCTCGG 2645

Qy       9426 CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGGGCACC 9485

Db       2646 CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGGGCACC 2705

Qy       9486 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTTGGTGA 9545

Db       2706 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTTGGTGA 2765

Qy       9546 GCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCCGGGGC 9605

Db       2766 GCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCCGGGGC 2825

Qy       9606 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT 9665

Db       2826 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT 2885

Qy       9666 GTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTCCTGAG 9725

Db       2886 GTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTCCTGAG 2945

Qy       9726 CTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGCTCCGG 9785

Db       2946 CTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGCTCCGG 3005

Qy       9786 GCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTAT 9845

Db       3006 GCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTCT 3065

Qy       9846 TGCATGAGGGGCGT 9859

Db       3066 TGGTGGATGTCAGT 3079
```

RESULT 11
AAV09024
ID AAV09024 standard; DNA; 3183 BP.
XX
AC AAV09024;
XX
DT 21-JUL-1998 (first entry)
XX
DE Homo sapiens 20q13 amplicon ZABC-1 cDNA sequence.
XX
KW 20q13 amplicon; chromosome 20; tumour; detection; ZABC-1 gene;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentation;
KW Leber's congenital amaurosis; zinc finger amplified in breast cancer; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .3183
FT /*tag= a
FT /product= "ZABC1 protein"
XX
PN WO9802539-A1.
XX
PD 22-JAN-1998.
XX
PF 15-JUL-1997; 97WO-US012343.
XX
PR 15-JAN-1996; 96US-00680395.
PR 16-OCT-1996; 96US-00731499.
PR 17-JAN-1997; 97US-00785532.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
XX
DR WPI; 1998-110587/10.
DR P-PSDB; AAW23975.
XX
PT New sequences from the 20q13 amplicon - used for detecting chromosomal
PT abnormalities, particularly tumours, and for developing products for
PT treating diseases.
XX
PS Claim 1; Page 67-69; 91pp; English.
XX
CC The sequence is that of the cDNA sequence encoding ZABC-1 (zinc finger
CC amplified in breast cancer). It maps to the core of the 20q13.2 amplicon
CC and is overexpressed in primary tumours and breast cancer cell lines
CC having 20q13.2 amplification. The sequence can be used as a probe for the
CC detection of chromosomal abnormalities at 20q13. It and other sequences
CC isolated from the 20q13 amplicon are consistently amplified in primary
CC tumours. These sequences are useful as probes or as probe targets for
CC monitoring the relative copy number of corresponding sequences from a
CC biological sample such as tumour cells. The sequences can also be used in
CC therapeutic applications for modulating the expression of the endogenous
CC gene or the activity of the gene product. Examples of therapeutic
CC approaches include antisense inhibition of gene expression, gene therapy,
CC and monoclonal antibodies that specifically bind the gene products. The
CC products can also be used in the treatment of other diseases, e.g. age-
CC related macular degeneration, Leber's congenital amaurosis and retinitis
CC pigmentation
XX

```
SQ   Sequence 3183 BP; 925 A; 796 C; 758 G; 698 T; 0 U; 6 Other;

  Query Match             15.5%;  Score 1554.8;  DB 2;  Length 3183;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1562;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       8286 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACATCTCT 8345
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1479 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACATCTCT 1538

Qy       8346 GAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 8405
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1539 GAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 1598

Qy       8406 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCGCAAAC 8465
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1599 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCGCAAAC 1658

Qy       8466 CAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCTGCAAA 8525
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1659 CAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCTGCAAA 1718

Qy       8526 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTCTCACC 8585
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1719 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTCTCACC 1778

Qy       8586 AGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGATAAAGT 8645
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1779 AGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGATAAAGT 1838

Qy       8646 GAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAC 8705
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1839 GAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAC 1898

Qy       8706 TCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGATGGCAG 8765
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1899 TCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGATGGCAG 1958

Qy       8766 TACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCTGACTG 8825
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1959 TACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGACACCGCAGCTGACTG 2018

Qy       8826 CAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCT 8885
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2019 CAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCT 2078

Qy       8886 TCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGTCCATT 8945
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2079 TCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGTCCATT 2138

Qy       8946 TTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAGCATAA 9005
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2139 TTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAGCATAA 2198

Qy       9006 ATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGACGTAC 9065
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2199 ATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGACGTAC 2258

Qy       9066 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGTAAACC 9125
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       2259 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCTCTAGTTTCTGTAAACC 2318

Qy       9126 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 9185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2319 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 2378

Qy       9186 GAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 9245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2379 GAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 2438

Qy       9246 CCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCCGCCAC 9305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2439 CCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCCGCCAC 2498

Qy       9306 CAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGATAAGAC 9365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2499 CAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGATAAGAC 2558

Qy       9366 AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACCCTCGG 9425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2559 AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACCCTCGG 2618

Qy       9426 CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGGGCACC 9485
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2619 CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGGGCACC 2678

Qy       9486 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTTGGTGA 9545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2679 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTTGGTGA 2738

Qy       9546 GCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCCGGGGC 9605
             |||||||||||||||||||||||| |||||||||||| |||||||||||||||||||||||
Db       2739 GCCCCTTCCAAAAAGACTGAAGTCCACCGTGGTTGCCCTTCACGTTGACCAGCCCGGGGC 2798

Qy       9606 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT 9665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2799 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT 2858

Qy       9666 GTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTCCTGAG 9725
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2859 GTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTCCTGAG 2918

Qy       9726 CTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGCTCCGG 9785
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2919 CTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGCTCCGG 2978

Qy       9786 GCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTAT 9845
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       2979 GCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTCT 3038

Qy       9846 TGCATGAGGGGCGT 9859
              ||   || |   ||
Db       3039 TGGTGGATGTCAGT 3052
```

```
RESULT 12
AAV04698
ID   AAV04698 standard; cDNA; 1507 BP.
XX
AC   AAV04698;
XX
DT   21-JUL-1998  (first entry)
XX
DE   Homo sapiens 20q13 amplicon cc49 transcript.
XX
KW   20q13 amplicon; chromosome 20; tumour; detection; C2H2 zinc finger;
KW   chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW   treatment; age-related macular degeneration; retinitis pigmentation;
KW   Leber's congenital amaurosis; ds.
XX
OS   Homo sapiens.
XX
PN   WO9802539-A1.
XX
PD   22-JAN-1998.
XX
PF   15-JUL-1997;   97WO-US012343.
XX
PR   15-JUL-1996;   96US-00680395.
PR   16-OCT-1996;   96US-00731499.
PR   17-JAN-1997;   97US-00785532.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Gray JW,  Collins CC,  Hwang S,  Godfrey T,  Kowbel D,  Rommens J;
XX
```

DR WPI; 1998-110587/10.
XX
PT New sequences from the 20q13 amplicon - used for detecting chromosomal
PT abnormalities, particularly tumours, and for developing products for
PT treating diseases.
XX
PS Claim 1; Page 61-62; 91pp; English.
XX
CC The sequence is that of a cDNA sequence cc49, which was isolated from the
CC 20q13 amplicon. It shows homology to C2H2 zinc finger genes. It can be
CC used as a probe for the detection of chromosomal abnormalities at 20q13.
CC It and other sequences isolated from the 20q13 amplicon are consistently
CC amplified in primary tumours. These sequences are useful as probes or as
CC probe targets for monitoring the relative copy number of corresponding
CC sequences from a biological sample such as tumour cells. The sequences
CC can also be used in therapeutic applications for modulating the
CC expression of the endogenous gene or the activity of the gene product.
CC Examples of therapeutic approaches include antisense inhibition of gene
CC expression, gene therapy, and monoclonal antibodies that specifically
CC bind the gene products. The products can also be used in the treatment of
CC other diseases, e.g. age-related macular degeneration, Leber's congenital
CC amaurosis and retinitis pigmentation
XX
SQ Sequence 1507 BP; 445 A; 354 C; 364 G; 340 T; 0 U; 4 Other;

Query Match 13.9%; Score 1387.2; DB 2; Length 1507;
Best Local Similarity 97.3%; Pred. No. 2.2e-294;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

```
Qy    2424 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGA 2483
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGC 61

Qy    2484 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGT 2543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGT 121

Qy    2544 TTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGG 2603
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 TTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGG 181

Qy    2604 ACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTTGGAACCCAGG 2663
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     182 ACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGG-TTTTGGAACCCAGG 240

Qy    2664 GCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTCCAGAAATTGC 2723
           ||||||||  |||||||||||| |||||||| |||||||||| ||||||||||||||||
Db     241 ACTGTTAGA-GGGTGAGTGACAGGTCTTAC-AGTGGCCTTAATCCAACTCCAGAAATTGC 298

Qy    2724 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 2783
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 358

Qy    2784 CTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAG 2843
           |||||||||||||||||| |||||  |||||||||||||||||||| ||||||||||||||
Db     359 CTCTTAATGTACATGGAT-GGCCAAGAGTGATTGGCAGCTCTCTTGCCAGTCCGATGGAG 417

Qy    2844 ATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGAA 2903
           |||||  ||||||||||||||||||||| || |||||||    | |           |||||
Db     418 ATGGA-GATGCCTTGTCAATGAAAGGGCCCNCTGTTGTCAATTCCGAGCTACACAAAGAA 476

Qy    2904 AAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 2962
           |||||||||  ||||  |||||||||| ||||||||||||||||||||||||||||||||
Db     477 AAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 536

Qy    2963 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 3022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     537 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 596

Qy    3023 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 3082
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     597 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 656

Qy    3083 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 3142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     657 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 716

Qy    3143 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 3202
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     717 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 776

Qy    3203 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 3262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     777 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 836

Qy    3263 TCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 3322
           |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db     837 TCACATGCGGACRCATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 896

Qy    3323 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTC 3382
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     897 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTC 956

Qy    3383 TCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 3442
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     957 TCCTTGCAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 1016

Qy    3443 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTC 3502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1017 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTC 1076

Qy    3503 TCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGACC 3562
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1077 TCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGACC 1136

Qy    3563 AAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGTT 3622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1137 AAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGTT 1196

Qy    3623 CACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTGCCAAGAAGT 3682
           |||||||||||||||||||:||||||||||:||||||||||||||||||||||||||||
Db    1197 CACCACCTTCCAGGCTTGGCAKCTGGCTACCAAAGGAAWAGTTGCCATTTGCCAAGAAGT 1256

Qy    3683 GAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGAGAAGGAGCT 3742
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1257 GAAGGAATTGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGAGAAGGAGCT 1316

Qy    3743 TGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACA 3802
           |||||||||||||||  |  ||||||||||||||||||||||||||||||||||||||||
Db    1317 TGGAGAAACAAATAAGAACCATTGTGCAGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACA 1376

Qy    3803 CTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAA 3862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1377 CTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAA 1436

Qy    3863 GCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCTGGTCTTGCA 3922
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1437 GCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCTGGTCTTGCA 1496

Qy    3923 CTCCAGGGTCC 3933
           ||||||||||| 
Db    1497 CTCCAGGGTCC 1507
```

RESULT 13
ADP03055/c
ID ADP03055 standard; cDNA; 124987 BP.
XX
AC ADP03055;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human housekeeping gene cDNA #96.
XX
KW ss; gene; human; housekeeping gene; capture probe; DNA microarray;
KW cancer.

```
XX
OS   Homo sapiens.
XX
PN   JP2004135552-A.
XX
PD   13-MAY-2004.
XX
PF   16-OCT-2002; 2002JP-00302278.
XX
PR   16-OCT-2002; 2002JP-00302278.
XX
PA   (NIGA ) NGK INSULATORS LTD.
XX
DR   WPI; 2004-404613/38.
XX
PT   Novel human housekeeping gene commonly expressed in different human
PT   tissues, useful as capture probe in DNA microarray.
XX
PS   Claim 1; SEQ ID NO 96; 940pp; Japanese.
XX
CC   The invention relates to a human housekeeping gene. The housekeeping gene
CC   is useful as a capture probe in DNA microarray. The housekeeping gene is
CC   useful as a reference with respect to the normal gene expression measured
CC   on various conditions, and an expression database of normal human tissue
CC   is useful as reference information for using a disease related gene for
CC   diagnosis or treatment. The housekeeping gene when used as capture probe,
CC   enables diagnosis of cancer. The present sequence represents a human
CC   housekeeping gene cDNA.
XX
SQ   Sequence 124987 BP; 31878 A; 32311 C; 31059 G; 29739 T; 0 U; 0 Other;

  Query Match             10.1%;  Score 1006.8;  DB 12;  Length 124987;
  Best Local Similarity   95.5%;  Pred. No. 5.9e-210;
  Matches 1083;  Conservative    0;  Mismatches   37;  Indels   14;  Gaps    4;

Qy          1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGG------AGACTTGCTCTGTTGCCCA 54
              ||||||||||||||| |||||| ||   ||| ||||      || || ||||||||||||||
Db       1128 CCATCATATTTCTTTTTTTTTGGGCAAGAGGGGGGAGTCAAAGTCTCGCTCTGTTGCCCA 1069

Qy         55 GGCTGGA--CCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTG 112
              |||||||  |||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1068 GGCTGGAATGCAGTGGCGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTG 1009

Qy        113 ATTCCCAAATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTT 172
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1008 ATTCCCAAGTAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTT 949

Qy        173 TAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTG 232
              |||||||||||||||||||||||||||| |||||||||||||||||||| ||||| |||
Db        948 TAGCAGATAAGGGGTTTCACCATGTTGGTCAGGCTGGTCTCCAACTCCTGACCTCAAGTG 889

Qy        233 ATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTC 292
              |||||||||||||||||||||||||||||||||||| |||||||||||||||| |||| |
Db        888 ATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTACAGGTGTGAGCCACTATGCCCGGC 829

Qy        293 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGT 352
              |||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db        828 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTGGAAAGGATGT 769

Qy        353 CAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAAT 412
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        768 CAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACATATTAGTCAAGTTTCTAAAAT 709

Qy        413 ACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTT 472
              |||||||||||||||||||| |||||||||||||| |     ||||| ||||||||||
Db        708 ACAGTAATAATGTTGAGAGCTGACGTAGGGACTAACTTAG-----TTTTTGTTTTTTTTT 654

Qy        473 TTTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAA 532
              |||||||||||||  |||||||||||||||||||||||||||| |||||||| ||||||
Db        653 TTTCAAATTCTCACCTAACTTTGATTTTGCTAAATAAGGACATT-AAAAAAAACCCAAAA 595

Qy        533 AACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCAT 592
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        594 AACTCCACTATTGCATATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCAT 535

Qy        593 CTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGT 652
              ||  |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        534 CTGCAAGTAGGAAGGGCCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGT 475

Qy        653 TGTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCC 712
              |||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 TGTCAGCATTTGTTAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCC 415

Qy        713 TGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATG 772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATG 355

Qy        773 TATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAAT 832
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        354 TATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAGT 295

Qy        833 GTGTAGCTCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTA 892
              |||||| |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        294 GTGTAGATCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTAGAGTTTA 235

Qy        893 ACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGT 952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 ACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGT 175

Qy        953 TGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGC 1012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 TGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGC 115

Qy       1013 CCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCT 1072
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        114 CCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAGAGACTTCAGGACATTCTTTCCT 55

Qy       1073 TTACCCTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATC 1126
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db          54 TTACCCTAATCCTTGATCTTCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATC 1

RESULT 14
ADS88553/c
ID   ADS88553 standard; cDNA; 124990 BP.
XX
AC   ADS88553;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human housekeeping gene cDNA sequence SEQ ID NO:96.
XX
KW   housekeeping gene; human; probe; hybridisation; DNA microarray;
KW   disease investigation; disease diagnosis; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004035785-A1.
XX
PD   29-APR-2004.
XX
PF   16-OCT-2002; 2002WO-JP010753.
XX
PR   16-OCT-2002; 2002WO-JP010753.
XX
PA   (NIGA ) NGK INSULATORS LTD.
XX
PI   Aburatani H,  Yamamoto S;
XX
DR   WPI; 2004-357219/33.
XX
PT   Human housekeeping and tissue-specific genes, gene sets and transcription
PT   products, probes and microarrays derived from them for reference use in
PT   investigation and diagnosis of disease.
XX
PS   Claim 5; SEQ ID NO 96; 372pp; Japanese.
```

```
XX
CC   The present invention describes 1189 housekeeping genes and gene sets
CC   containing them, expressed in 35 different human tissues. Also described:
CC   (1) RNA and cDNA transcription products of the housekeeping genes; (2)
CC   oligonucleotide probes hybridising to the housekeeping genes; (3) DNA
CC   microarrays containing the oligonucleotide probes; (4) 1704 genes
CC   expressed specifically in particular tissues; (5) RNA and cDNA
CC   transcription products of the tissue-specific genes; (6) oligonucleotide
CC   probes hybridising to the tissue-specific genes; and (7) DNA microarrays
CC   containing the oligonucleotide probes. The tissue-specific genes are
CC   expressed specifically in one of brain, thymus, pituitary, thyroid,
CC   trachea, lung, breast, skin, skeletal muscle, heart, liver, spleen,
CC   kidney, adrenal gland, pancreas, stomach, small intestine, colon,
CC   bladder, prostate gland, testis, ovary, placenta, uterus, bone marrow,
CC   foetal brain or foetal liver. The housekeeping genes can be used as
CC   reference genes for the investigation and diagnosis of disease. The
CC   present sequence represents a human housekeeping gene cDNA sequence which
CC   is given in the exemplification of the present invention.
XX
SQ   Sequence 124990 BP; 31879 A; 32311 C; 31060 G; 29740 T; 0 U; 0 Other;

  Query Match             10.1%;  Score 1006.8;  DB 13;  Length 124990;
  Best Local Similarity   95.5%;  Pred. No. 5.9e-210;
  Matches 1083;  Conservative    0;  Mismatches   37;  Indels   14;  Gaps    4;

Qy          1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGG------AGACTTGCTCTGTTGCCCA 54
              ||||||||||||||| ||||||  ||   ||| ||||      || || ||||||||||||
Db       1128 CCATCATATTTCTTTTTTTTTGGGCAAGAGGGGGGAGTCAAAGTCTCGCTCTGTTGCCCA 1069

Qy         55 GGCTGGA--CCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTG 112
              |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1068 GGCTGGAATGCAGTGGCGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTG 1009

Qy        113 ATTCCCAAATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTT 172
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1008 ATTCCCAAGTAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTT 949

Qy        173 TAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTG 232
              |||||||||||||||||||||||||||||| |||||||||||||||||||| ||||| |||
Db        948 TAGCAGATAAGGGGTTTCACCATGTTGGTCAGGCTGGTCTCCAACTCCTGACCTCAAGTG 889

Qy        233 ATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTC 292
              |||||||||||||||||||||||||||||||||||| |||||||||||||||| |||| |
Db        888 ATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTACAGGTGTGAGCCACTATGCCCGGC 829

Qy        293 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGT 352
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        828 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTGGAAAGGATGT 769

Qy        353 CAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAAT 412
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        768 CAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACATATTAGTCAAGTTTCTAAAAT 709

Qy        413 ACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTT 472
              |||||||||||||||||||||| |||||||||||| |     ||||| |||||||||
Db        708 ACAGTAATAATGTTGAGAGCTGACGTAGGGACTAACTTAG-----TTTTTGTTTTTTTTT 654

Qy        473 TTTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAACCAAAA 532
              |||||||||||||  |||||||||||||||||||||||||||||| |||||||| ||||||
Db        653 TTTCAAATTCTCACCTAACTTTGATTTTGCTAAATAAGGACATT-AAAAAAAACCCAAAA 595

Qy        533 AACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCAT 592
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        594 AACTCCACTATTGCATATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCAT 535

Qy        593 CTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGT 652
              ||  |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        534 CTGCAAGTAGGAAGGGCCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGT 475

Qy        653 TGTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCC 712
              |||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 TGTCAGCATTTGTTAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCC 415

Qy        713 TGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATG 772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATG 355

Qy        773 TATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAAT 832
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        354 TATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAGT 295

Qy        833 GTGTAGCTCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTA 892
              |||||| ||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        294 GTGTAGATCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTAGAGTTTA 235

Qy        893 ACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGT 952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 ACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGT 175

Qy        953 TGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGC 1012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 TGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGC 115

Qy       1013 CCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCT 1072
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        114 CCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAGAGACTTCAGGACATTCTTTCCT 55

Qy       1073 TTACCCTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATC 1126
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db          54 TTACCCTAATCCTTGATCTTCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATC 1


RESULT 15
ACN44984
ID   ACN44984 standard; DNA; 26345 BP.
XX
AC   ACN44984;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Mouse genomic sequence mCG6549.
XX
KW   Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX
OS   Mus musculus.
XX
PN   WO2003073826-A2.
XX
PD   12-SEP-2003.
XX
PF   28-FEB-2003; 2003WO-US006235.
XX
PR   01-MAR-2002; 2002US-00087192.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
XX
DR   WPI; 2003-328604/31.
XX
PT   Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT   comprises a nucleotide sequence.
XX
PS   Claim 1; SEQ ID NO 1705; 0pp; English.
XX
CC   The present invention relates to novel DNA and protein sequences which
CC   are associated with carcinomas. The sequences are useful for: (i) for
CC   screening drug candidates; (ii) for screening of bioactive agent capable
CC   of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC   a bioactive agent capable of modulating the activity of CAP; (iv) for
CC   evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC   carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC   carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC   (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC   determining Carcinoma Associated (CA) gene copy number. In addition, the
CC   CA genes are useful as DNA vaccines and the CAP are useful as markers of
```

```
CC   carcinoma including lymphoma. The present sequence is one such CA coding
CC   sequence. Note: This patent is an equivalent to basic patent
CC   US2002182586A1, for which no sequence data was published
XX
SQ   Sequence 26345 BP; 6286 A; 5855 C; 6718 G; 7486 T; 0 U; 0 Other;

  Query Match             8.2%;  Score 820;  DB 11;  Length 26345;
  Best Local Similarity   68.1%;  Pred. No. 3.5e-169;
  Matches 1391;  Conservative    0;  Mismatches  550;  Indels  101;  Gaps   14;

Qy       2132 AACTTGACATTTTAGATTTAAGTCGGTAAAGCACTGATTTAAACTGGATTTTAACTGGAT 2191
Db       9394 AAATTTGGCATTTATATTTCAGTGAGCAGAGCTCTGCCTTAAGTCTCCTGACTGGTTGAA 9453

Qy       2192 GAAATTCTGATTTAATAAGTGTACTGACTGGATAAAATGCCAATGATTTAATTAACAAGC 2251
Db       9454 ATTCTACTCACTTAGTAAGTGTATCCACTGGGAAATATGCTAATGATTTAATTAACAAAC 9513

Qy       2252 ACGTTTAACAGGATGCCCTATATATTAGTTAAAAGTGAAGCAATTGAATTAGGTACCTTC 2311
Db       9514 ATGTTTAGCAGGCTGCATTATGGATTAGTTAAAGGTGAGGCAGTTGACTCTAGTCCCTTT 9573

Qy       2312 TCTGCTGCGTGGAAAAGACCGTATGACTCACCCACACCAGCCTTCTCTTCGCTCTGAGTG 2371
Db       9574 T-------GTGGAGAAGATCATGTGACTCATAGGCCAAGCCCTGTTCCTGAGCCTTCCTC 9626

Qy       2372 TAGCTAACCGTTTCTGTTTTTTTTCCTCTAGGGTTTGGAAATCCCTTGTCTCCAGGTTGC 2431
Db       9627 ACTGAGAC-----TTGTTTTTTTCCTTTCAGATTTTGGAAATCCCTTGTCTTCACATTGC 9681

Qy       2432 TGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGAACTAATG- 2490
Db       9682 CGGGATCAACGTCTTGCTCAACTGAGCCACTCA-CCAGTGGAGATGGACAGAGCTGTTGC 9740

Qy       2491 CTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGTTTGTGGA 2550
Db       9741 CCTCGTGCTGACGCACATCTGACTCACAGCCTGGGGCAGCTTGGATGCCTTGTTTGTGCA 9800

Qy       2551 AAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGGACTATAC 2610
Db       9801 AAGAGCCA---TCTCCATCACTGAGCTCTCCCAAGCTCC-----------GGAATATAC 9845

Qy       2611 ACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTTGGAACCCAGGGCTGTTA 2670
Db       9846 TCGTCCCTCTGACCTT---------AGAGGTGCTCTGATTTTGACACCCAGGGCTCTTAG 9896

Qy       2671 GAGGGGTGAGTGAC-AAGTCTTACAAGTGGCCTTATTCCAACTCCAGAAATTGCC----- 2724
Db       9897 AGGTTCTGAGTGGTGATATCTCACAGGCGGCCTTATTCTAACTGCAGAGATCCTCAGCAG 9956

Qy       2725 CAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCC 2784
Db       9957 ATAGTGGATTACCCATCCTGTCAGACCCTGAAGTGGTTGGCAGCATGCCGACCCAGTCCC 10016

Qy       2785 TCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGA 2844
Db       10017 TCCTCGTGTACATGGACGGGCCGGAAGTCCTCAGCAGCTCTCTAGGCTCCCAGATGGAGG 10076

Qy       2845 TGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGAAA 2904
Db       10077 TGGATGATGCTGTGCCCATAAAAGGGCCGGTGGCAGTCCCCTTCCGAGCTGCTCAGGAGA 10136

Qy       2905 AAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGACCT 2964
Db       10137 AGAGCATGGCCGTGGCAGAGGGCCACATGCCCCTGGATTGCATGTTCTGCAGCCAGGTCT 10196

Qy       2965 TCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTGTG 3024
Db       10197 TCTCTCAGGCGGAGGATCTCAGTCAGCACGTGCTGCTGCAGCACCGGCCCACCCTCTGCG 10256

Qy       3025 AACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGC 3084
Db       10257 AGCCAGCTGTCCTGCGTGTGGAGGCCGAGTACCTAAGTCCCCTTGATAAAGCTCTGGAGC 10316

Qy       3085 GAACAGAACCTCCCAAGGAAAAGAATTGCAAGGA---AAATGAATTTAGCTGTGAGGTAT 3141
Db       10317 CAACAGAGCCAGCATTGGAGAAGAGTGGCGAAGACCCCGAGGAGTTGAGCTGTGATGTGT 10376

Qy       3142 GTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATT 3201
Db       10377 GTGGGCAGACATTCCCAGTGGCTTTTGATGTTGAGAGCCACATGAAGAAGCATAAGGACT 10436

Qy       3202 CTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAA 3261
Db       10437 CCTTCACGTATGGGTGCAGCATGTGCGGGAGGAGATTCAAGGAGCCGTGGTTCCTGAAGA 10496

Qy       3262 ATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGG 3321
Db       10497 ACCACATGCGGACACACAATGGCAAGTCTGGCACCAGGAGCAAGCTTCAGCAAGGCATGG 10556

Qy       3322 AGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCT 3381
Db       10557 AG---AGTCCAGTCACCATCAATGAAGTGGTCCAGCCGCACGCCCCTGGGAGCATCTCCA 10613

Qy       3382 CTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTG 3441
Db       10614 CGCCCTACAAGATCTGCATGGTCTGCGGCTTCCTCTTCCCAAATAAGCAGAGCCTCATTG 10673

Qy       3442 AGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACT 3501
Db       10674 AGCACAGCAAGGTTCACGCCAAAGAAACTGTCCCCAGTGCCAGCAACGTTGCCCCTGATG 10733

Qy       3502 CTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGAC 3561
Db       10734 ACCACCGAGAGGAACCCACGTCCCCGAGGGAAGAGTTGCTGCAGTTTTTGAACTTGAGAC 10793

Qy       3562 CAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGT 3621
Db       10794 CCAGATCAACTGCAGGTAGTACAGTGAAGCCCATGACCTGCATACCTCAGCTTGACCCGT 10853

Qy       3622 TCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTGCC---AAG 3678
Db       10854 TCACCACCTACCAGGCATGGCAGTTGGCTACCAAAGGAAAAGTGGCCGTTGCCCAGGAAG 10913

Qy       3679 AAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGAGAAGG 3738
Db       10914 AGGTGAAAGAGTCAGGCCAAGAAGGAAGCACAGACAATGACGACTCATGCTCAGAGAAAG 10973

Qy       3739 AGCTTGGAG----------------------------------------------AAACAA 3753
Db       10974 AGGAACTAGGGGAAATATGGGTTGGGGGTAAGGCGGAAGGGTCTGGAAAGTCCAAAACAA 11033

Qy       3754 ATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACACTCCCACGGCG 3813
Db       11034 GTAAAAGCAGTTGTCCAGGTCTCTCCCAAGACAAAGAGAAGCCTAGACATGCTAATAGTG 11093

Qy       3814 AAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAAGCCCACTCACT 3873
Db       11094 AAGTGCCTTCTGGGGATAGTGACCCCAAGTTGTCCAGTAGCAAGGAGAAGCCCACGCACT 11153

Qy       3874 GCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCTGGTCTTGCACTCCAGGGTCC 3933
Db       11154 GTTCTGAGTGCAGCAAAGCCTTCAGGACATACCACCAGCTCGTCCTGCACTCGAGGGTGC 11213

Qy       3934 ACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCATGTCTGTGGACGGGAGGCAGC 3993
Db       11214 ACAGGAAGGACAGGAGGACTGATGCCCTGTCACCCACCATGGCTGTGGATGCAAGGCAGC 11273

Qy       3994 CGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGAAAATGGAGCCGTGGATCGAG 4053
Db       11274 CTGGGACCTGCTCCCCAGACCTCAGCACCACTCTGGAAGACAGTGGGGCCGGGGACCGA- 11332

Qy       4054 GGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGGTA 4113
Db       11333 --GAAGGGGGCTCTGAAGACGGGTCTGAGGATGGACTCCCTGACGGGCTCCATTTGGGTG 11390
```

```
Qy       4114 AG 4115
             ||
Db      11391 AG 11392
```

Search completed: July 31, 2005, 12:46:01
Job time : 3149 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 10:56:42 ; Search time 27398 Seconds
(without alignments)
17685.676 Million cell updates/sec

Title: US-08-731-499-9_COPY_1_10000
Perfect score: 10000
Sequence: 1 CCATCATATTTCTTATTTTT.....ATTCTTAGATACGGCAGTGG 10000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

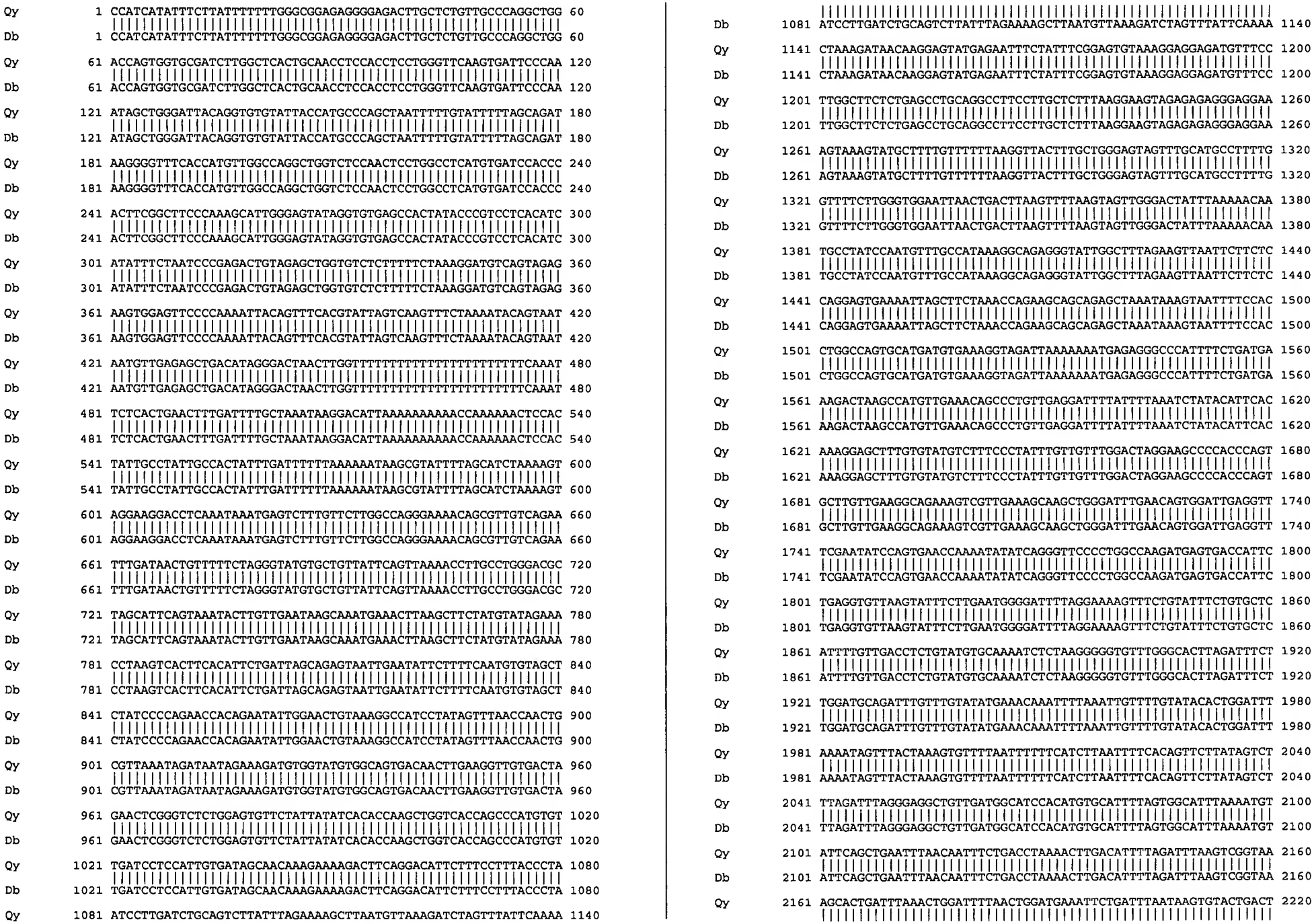
SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	10000	100.0	10365	6	BD085733	BD085733 Genes fro
c 2	10000	100.0	121143	9	AF312915	AF312915 Homo sapi
c 3	9980.6	99.8	128871	9	AL157838	AL157838 Human DNA
4	9943.6	99.4	105023	2	AC116668	AC116668 Trypanoso
c 5	2132	21.3	165471	2	AC015713	AC015713 Homo sapi
6	2015	20.2	111123	9	AP000365	AP000365 Homo sapi
c 7	2015	20.2	187590	2	AC141274	AC141274 Homo sapi
8	1639.4	16.4	5632	6	AR157086	AR157086 Sequence
9	1639.4	16.4	5632	9	AF041259	AF041259 Homo sapi
10	1639.4	16.4	5633	6	CQ726295	CQ726295 Sequence
11	1559.6	15.6	3186	6	BD085734	BD085734 Genes fro
12	1387.2	13.9	1507	6	AR070326	AR070326 Sequence
13	1387.2	13.9	1507	6	BD085727	BD085727 Genes fro
c 14	1223.6	12.2	59586	9	HS449O17	Z98749 Human DNA s
c 15	1006.8	10.1	124990	9	HS434P1	Z97056 Human DNA s
16	850.6	8.5	246332	2	AC099079	AC099079 Rattus no
17	828.8	8.3	184420	2	AC118373	AC118373 Rattus no
c 18	820	8.2	114285	10	AL844576	AL844576 Mouse DNA
19	820	8.2	208979	2	AC023610	AC023610 Mus muscu
20	345	3.5	46985	2	AC020762	AC020762 Homo sapi
21	335.8	3.4	469	6	BD060160	BD060160 Secreted
22	274.8	2.7	1196	10	BC046393	BC046393 Mus muscu
c 23	244.6	2.4	251	11	G65441	G65441 stdJ434P1_2
24	239.8	2.4	191540	9	AC144780	AC144780 Pan trogl
c 25	235.8	2.4	103287	2	AC084034	AC084034 Homo sapi
c 26	234	2.3	110000	2	BX572642_0	BX572642 Homo sapi
c 27	234	2.3	113051	9	AL353150	AL353150 Human DNA
c 28	233.8	2.3	172571	9	AC006064	AC006064 Homo sapi
c 29	233.4	2.3	83639	9	AL353787	AL353787 Human DNA
c 30	232.6	2.3	68620	2	AC031999	AC031999 Homo sapi
31	232.6	2.3	118695	9	HSJ672M15	AL049643 Human DNA
32	232	2.3	3512	5	BC072191	BC072191 Xenopus l
c 33	232	2.3	31484	9	AC131958	AC131958 Homo sapi
34	232	2.3	56330	2	AL353694_3	Continuation (4 of
35	231.2	2.3	177964	9	AC016705	AC016705 Homo sapi
36	231.2	2.3	198707	9	AC108451	AC108451 Homo sapi
37	231	2.3	156839	9	AC107948	AC107948 Homo sapi
c 38	231	2.3	166973	9	AC084117	AC084117 Homo sapi
39	230.8	2.3	110000	2	AL831785_2	Continuation (3 of
40	230.8	2.3	131682	9	AL672277	AL672277 Human DNA
c 41	230.4	2.3	2081	9	HSM806272	BX537592 Homo sapi
42	230.2	2.3	172579	9	AC008064	AC008064 Homo sapi
c 43	229.8	2.3	48593	2	AC113138	AC113138 Homo sapi
c 44	229.8	2.3	218892	9	AC007225	AC007225 Homo sapi
45	229.6	2.3	71819	9	AC007536	AC007536 Homo sapi

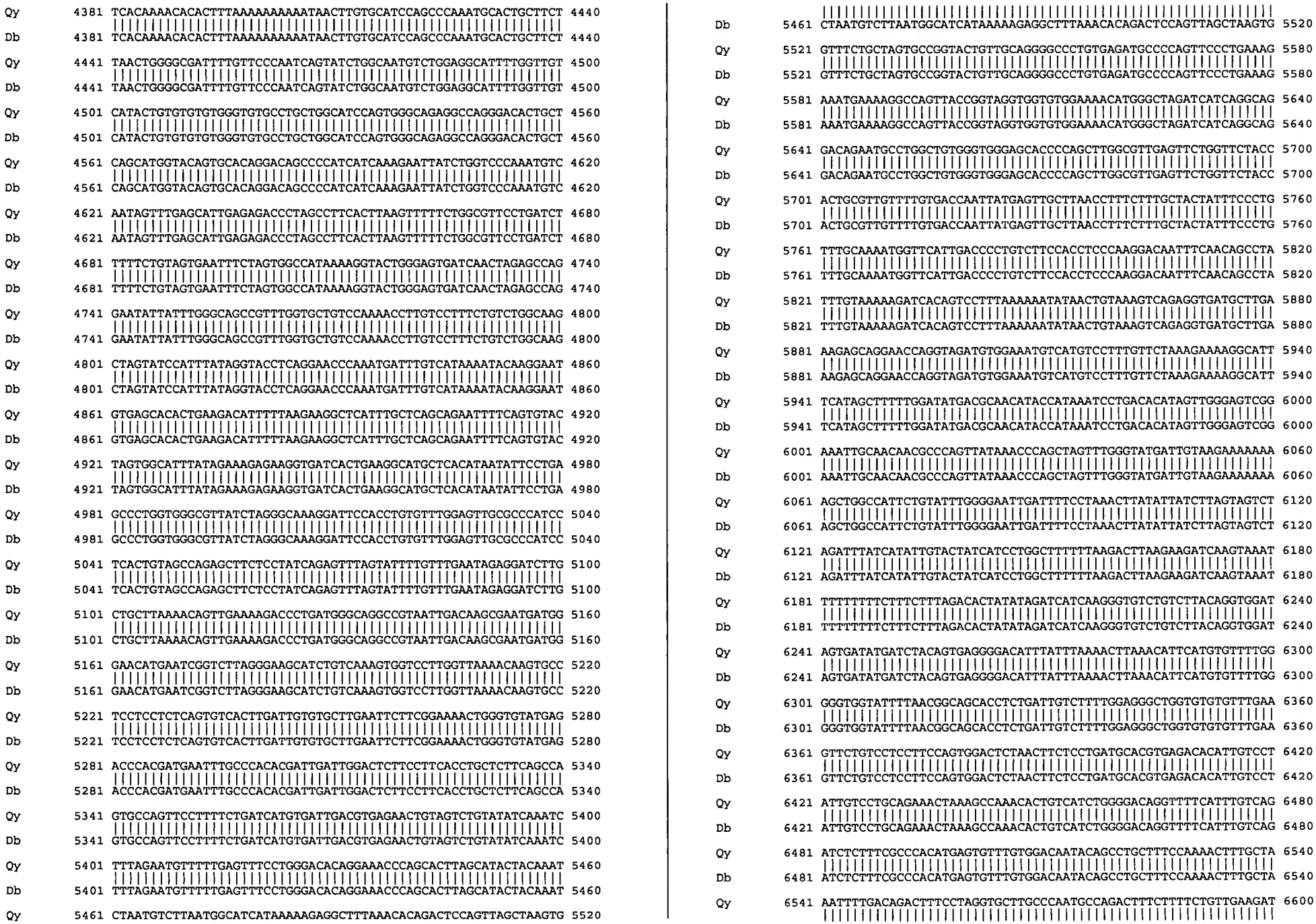
ALIGNMENTS

RESULT 1
BD085733
LOCUS BD085733 10365 bp DNA linear PAT 27-AUG-2002
DEFINITION Genes from the 20q13 amplicon and their uses.
ACCESSION BD085733
VERSION BD085733.1 GI:22631343
KEYWORDS JP 2001524802-A/9.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 10365)
AUTHORS Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and Rommens,J.
TITLE Genes from the 20q13 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 9 04-DEC-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
COMMENT OS Artificial Sequence
PN JP 2001524802-A/9
PD 04-DEC-2001
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR 17-JAN-1997 US 08/785532
PI JOE W GRAY,COLIN CONRAD COLLINS,SOO IN HWANG,TONY GODFREY, PI DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11,C12Q1/68,A61K48/00
CC Description of Artificial Sequence:Genomic Sequence encoding
CC ZABC1
FH Key Location/Qualifiers
FT source 1. .10365
FT /organism='Artificial Sequence'.
FEATURES Location/Qualifiers
source 1. .10365
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN

Query Match 100.0%; Score 10000; DB 6; Length 10365;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 10000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


Qy 1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 60
Db 1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 60
Qy 61 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 120
Db 61 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 120
Qy 121 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 180
Db 121 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 180
Qy 181 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 240
Db 181 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 240
Qy 241 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300
Db 241 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300
Qy 301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 360
Db 301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 360
Qy 361 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 420
Db 361 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 420
Qy 421 AATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTTCAAAT 480
Db 421 AATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTTCAAAT 480
Qy 481 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAACCAAAAAACTCCAC 540
Db 481 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAACCAAAAAACTCCAC 540
Qy 541 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 600
Db 541 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 600
Qy 601 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 660
Db 601 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 660
Qy 661 TTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 720
Db 661 TTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 720
Qy 721 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 780
Db 721 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 780
Qy 781 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 840
Db 781 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 840
Qy 841 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 900
Db 841 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 900
Qy 901 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 960
Db 901 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 960
Qy 961 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 1020
Db 961 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 1020
Qy 1021 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 1080
Db 1021 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 1080
Qy 1081 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 1140
Db 1081 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 1140
Qy 1141 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 1200
Db 1141 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 1200
Qy 1201 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 1260
Db 1201 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 1260
Qy 1261 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 1320
Db 1261 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 1320
Qy 1321 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 1380
Db 1321 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 1380
Qy 1381 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 1440
Db 1381 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 1440
Qy 1441 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 1500
Db 1441 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 1500
Qy 1501 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATGA 1560
Db 1501 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATGA 1560
Qy 1561 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 1620
Db 1561 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 1620
Qy 1621 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 1680
Db 1621 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 1680
Qy 1681 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 1740
Db 1681 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 1740
Qy 1741 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 1800
Db 1741 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 1800
Qy 1801 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
Db 1801 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
Qy 1861 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 1920
Db 1861 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 1920
Qy 1921 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 1980
Db 1921 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 1980
Qy 1981 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 2040
Db 1981 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 2040
Qy 2041 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 2100
Db 2041 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 2100
Qy 2101 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 2160
Db 2101 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 2160
Qy 2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 2220

```
Db       2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 2220
Qy       2221 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 2280

Qy       2281 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 2340

Qy       2341 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTCT 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTCT 2400

Qy       2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 2460

Qy       2461 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520

Qy       2521 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2580

Qy       2581 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 2640

Qy       2641 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2700

Qy       2701 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2760

Qy       2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2820

Qy       2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2880

Qy       2881 TTCCATTCCGAGCTACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 2940

Qy       2941 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 3000

Qy       3001 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 3060

Qy       3061 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 3120

Qy       3121 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 3180

Qy       3181 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 3240

Qy       3241 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 3300

Qy       3301 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 3360

Qy       3361 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 3420

Qy       3421 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 3480

Qy       3481 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 3540

Qy       3541 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 3600

Qy       3601 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 3660

Qy       3661 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 3720

Qy       3721 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 3780

Qy       3781 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 3840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 3840

Qy       3841 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 3900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 3900

Qy       3901 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 3960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 3960

Qy       3961 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 4020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 4020

Qy       4021 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 4080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 4080

Qy       4081 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 4140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 4140

Qy       4141 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 4200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 4200

Qy       4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 4260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 4260

Qy       4261 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 4320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 4320

Qy       4321 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 4380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4321 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 4380
```


Qy 4381 TCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 4440
Db 4381 TCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 4440
Qy 4441 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 4500
Db 4441 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 4500
Qy 4501 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 4560
Db 4501 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 4560
Qy 4561 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 4620
Db 4561 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 4620
Qy 4621 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 4680
Db 4621 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 4680
Qy 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATCAACTAGAGCCAG 4740
Db 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATCAACTAGAGCCAG 4740
Qy 4741 GAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGCAAG 4800
Db 4741 GAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGCAAG 4800
Qy 4801 CTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGGAAT 4860
Db 4801 CTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGGAAT 4860
Qy 4861 GTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTGTAC 4920
Db 4861 GTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTGTAC 4920
Qy 4921 TAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCCTGA 4980
Db 4921 TAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCCTGA 4980
Qy 4981 GCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCATCC 5040
Db 4981 GCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCATCC 5040
Qy 5041 TCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATCTTG 5100
Db 5041 TCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATCTTG 5100
Qy 5101 CTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGATGG 5160
Db 5101 CTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGATGG 5160
Qy 5161 GAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGTGCC 5220
Db 5161 GAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGTGCC 5220
Qy 5221 TCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTATGAG 5280
Db 5221 TCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTATGAG 5280
Qy 5281 ACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAGCCA 5340
Db 5281 ACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAGCCA 5340
Qy 5341 GTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAAATC 5400
Db 5341 GTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAAATC 5400
Qy 5401 TTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAAT 5460
Db 5401 TTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAAT 5460
Qy 5461 CTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAAGTG 5520
Db 5461 CTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAAGTG 5520
Qy 5521 GTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGAAAG 5580
Db 5521 GTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGAAAG 5580
Qy 5581 AAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGGCAG 5640
Db 5581 AAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGGCAG 5640
Qy 5641 GACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCTACC 5700
Db 5641 GACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCTACC 5700
Qy 5701 ACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCCCTG 5760
Db 5701 ACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCCCTG 5760
Qy 5761 TTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGCCTA 5820
Db 5761 TTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGCCTA 5820
Qy 5821 TTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCTTGA 5880
Db 5821 TTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCTTGA 5880
Qy 5881 AAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGCATT 5940
Db 5881 AAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGCATT 5940
Qy 5941 TCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGTCGG 6000
Db 5941 TCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGTCGG 6000
Qy 6001 AAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAAAAA 6060
Db 6001 AAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAAAAA 6060
Qy 6061 AGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAGTCT 6120
Db 6061 AGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAGTCT 6120
Qy 6121 AGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTAAAT 6180
Db 6121 AGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTAAAT 6180
Qy 6181 TTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTGGAT 6240
Db 6181 TTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTGGAT 6240
Qy 6241 AGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTTTGG 6300
Db 6241 AGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTTTGG 6300
Qy 6301 GGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTTGAA 6360
Db 6301 GGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTTGAA 6360
Qy 6361 GTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGTCCT 6420
Db 6361 GTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGTCCT 6420
Qy 6421 ATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTGTCAG 6480
Db 6421 ATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTGTCAG 6480
Qy 6481 ATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTGCTA 6540
Db 6481 ATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTGCTA 6540
Qy 6541 AATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAAGAT 6600

```
Db      6541 AATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAAGAT 6600
Qy      6601 TAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCTTAT 6660
Db      6601 TAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCTTAT 6660
Qy      6661 TTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTATTC 6720
Db      6661 TTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTATTC 6720
Qy      6721 TGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCTGCC 6780
Db      6721 TGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCTGCC 6780
Qy      6781 TGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATCTTT 6840
Db      6781 TGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATCTTT 6840
Qy      6841 CTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCAGAA 6900
Db      6841 CTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCAGAA 6900
Qy      6901 AAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGCATT 6960
Db      6901 AAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGCATT 6960
Qy      6961 TGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCATGCT 7020
Db      6961 TGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCATGCT 7020
Qy      7021 TTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTCAGC 7080
Db      7021 TTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTCAGC 7080
Qy      7081 TACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGAAAA 7140
Db      7081 TACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGAAAA 7140
Qy      7141 ATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCAAAT 7200
Db      7141 ATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCAAAT 7200
Qy      7201 TATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAACCA 7260
Db      7201 TATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAACCA 7260
Qy      7261 TGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTCTGC 7320
Db      7261 TGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTCTGC 7320
Qy      7321 CAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATTTTT 7380
Db      7321 CAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATTTTT 7380
Qy      7381 ATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTTTCT 7440
Db      7381 ATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTTTCT 7440
Qy      7441 TTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTGGTG 7500
Db      7441 TTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTGGTG 7500
Qy      7501 CAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAGCCT 7560
Db      7501 CAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAGCCT 7560
Qy      7561 CCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTTTAG 7620
Db      7561 CCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTTTAG 7620
Qy      7621 TAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATC 7680
Db      7621 TAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATC 7680
```

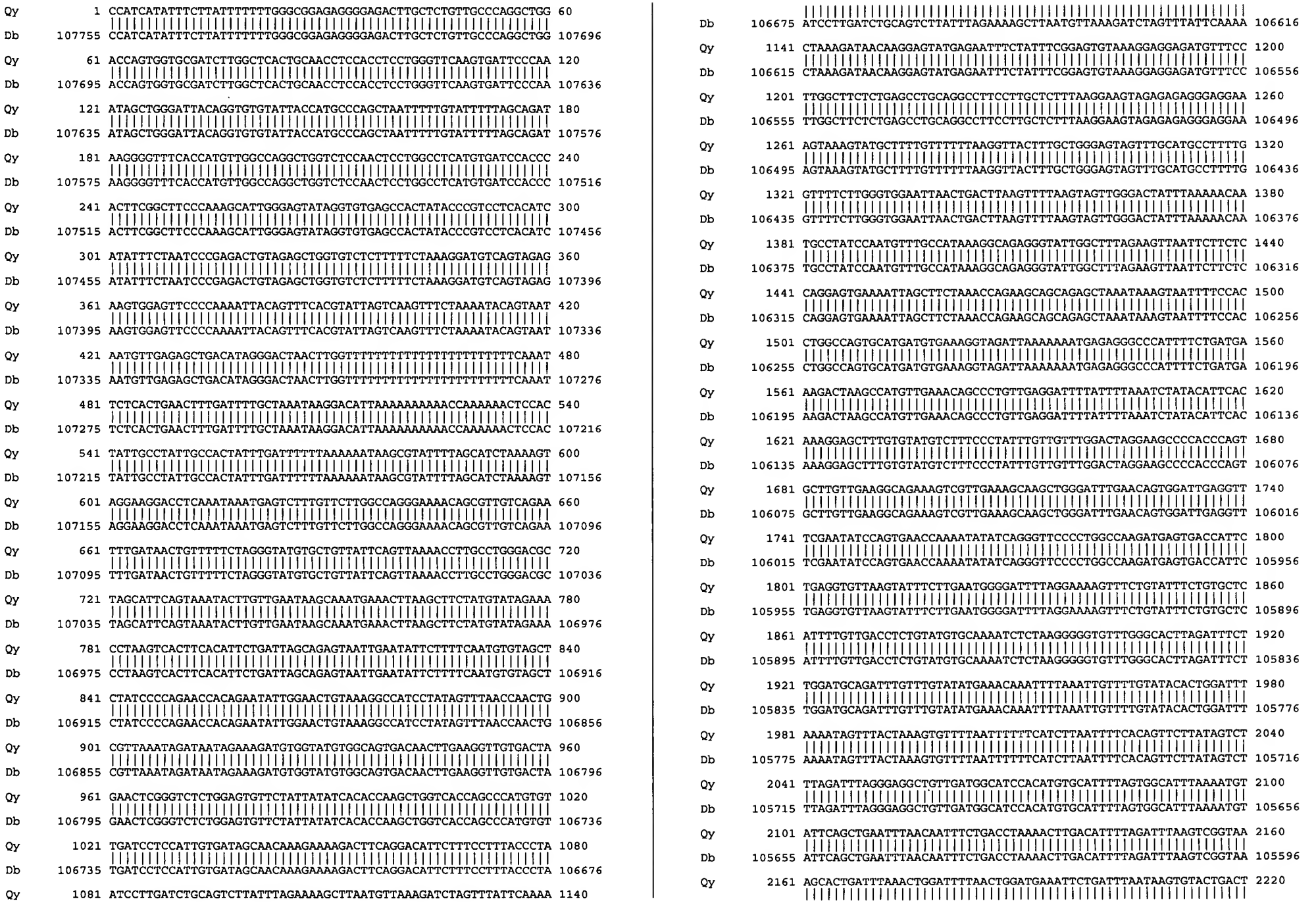
```
Qy      7681 CGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTGAGC 7740
Db      7681 CGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTGAGC 7740
Qy      7741 CACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTTATC 7800
Db      7741 CACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTTATC 7800
Qy      7801 TTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAATGAA 7860
Db      7801 TTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAATGAA 7860
Qy      7861 ATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTAGGA 7920
Db      7861 ATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTAGGA 7920
Qy      7921 GGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCAAGT 7980
Db      7921 GGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCAAGT 7980
Qy      7981 TATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATATTTC 8040
Db      7981 TATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATATTTC 8040
Qy      8041 CCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATTGAA 8100
Db      8041 CCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATTGAA 8100
Qy      8101 GATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGAGAA 8160
Db      8101 GATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGAGAA 8160
Qy      8161 ATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAATAA 8220
Db      8161 ATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAATAA 8220
Qy      8221 ATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTTCTT 8280
Db      8221 ATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTTCTT 8280
Qy      8281 GCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACA 8340
Db      8281 GCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACA 8340
Qy      8341 TCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAA 8400
Db      8341 TCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAA 8400
Qy      8401 GTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCG 8460
Db      8401 GTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCG 8460
Qy      8461 CAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCT 8520
Db      8461 CAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCT 8520
Qy      8521 GCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTC 8580
Db      8521 GCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTC 8580
Qy      8581 TCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGAT 8640
Db      8581 TCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGAT 8640
Qy      8641 AAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTT 8700
Db      8641 AAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTT 8700
Qy      8701 GAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGAT 8760
Db      8701 GAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGAT 8760
```

```
Qy       8761 GGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCT 8820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8761 GGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCT 8820

Qy       8821 GACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGG 8880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8821 GACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGG 8880

Qy       8881 GCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGT 8940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8881 GCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGT 8940

Qy       8941 CCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAG 9000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8941 CCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAG 9000

Qy       9001 CATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGA 9060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9001 CATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGA 9060

Qy       9061 CGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGT 9120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9061 CGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGT 9120

Qy       9121 AAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGG 9180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9121 AAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGG 9180

Qy       9181 AAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACT 9240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9181 AAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACT 9240

Qy       9241 TTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCC 9300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9241 TTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCC 9300

Qy       9301 GCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGAT 9360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9301 GCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGAT 9360

Qy       9361 AAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACC 9420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9361 AAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACC 9420

Qy       9421 CTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGG 9480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9421 CTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGG 9480

Qy       9481 GCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTT 9540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9481 GCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTT 9540

Qy       9541 GGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCC 9600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9541 GGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCC 9600

Qy       9601 GGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACA 9660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9601 GGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACA 9660

Qy       9661 TCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTC 9720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9661 TCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTC 9720

Qy       9721 CTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGC 9780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9721 CTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGC 9780

Qy       9781 TCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAA 9840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9781 TCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAA 9840

Qy       9841 GGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAATCCA 9900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9841 GGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAATCCA 9900

Qy       9901 GGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGTTCT 9960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9901 GGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGTTCT 9960

Qy       9961 TGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000
             |||||||||||||||||||||||||||||||||||||||||
Db       9961 TGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000
```

RESULT 2
AF312915/c

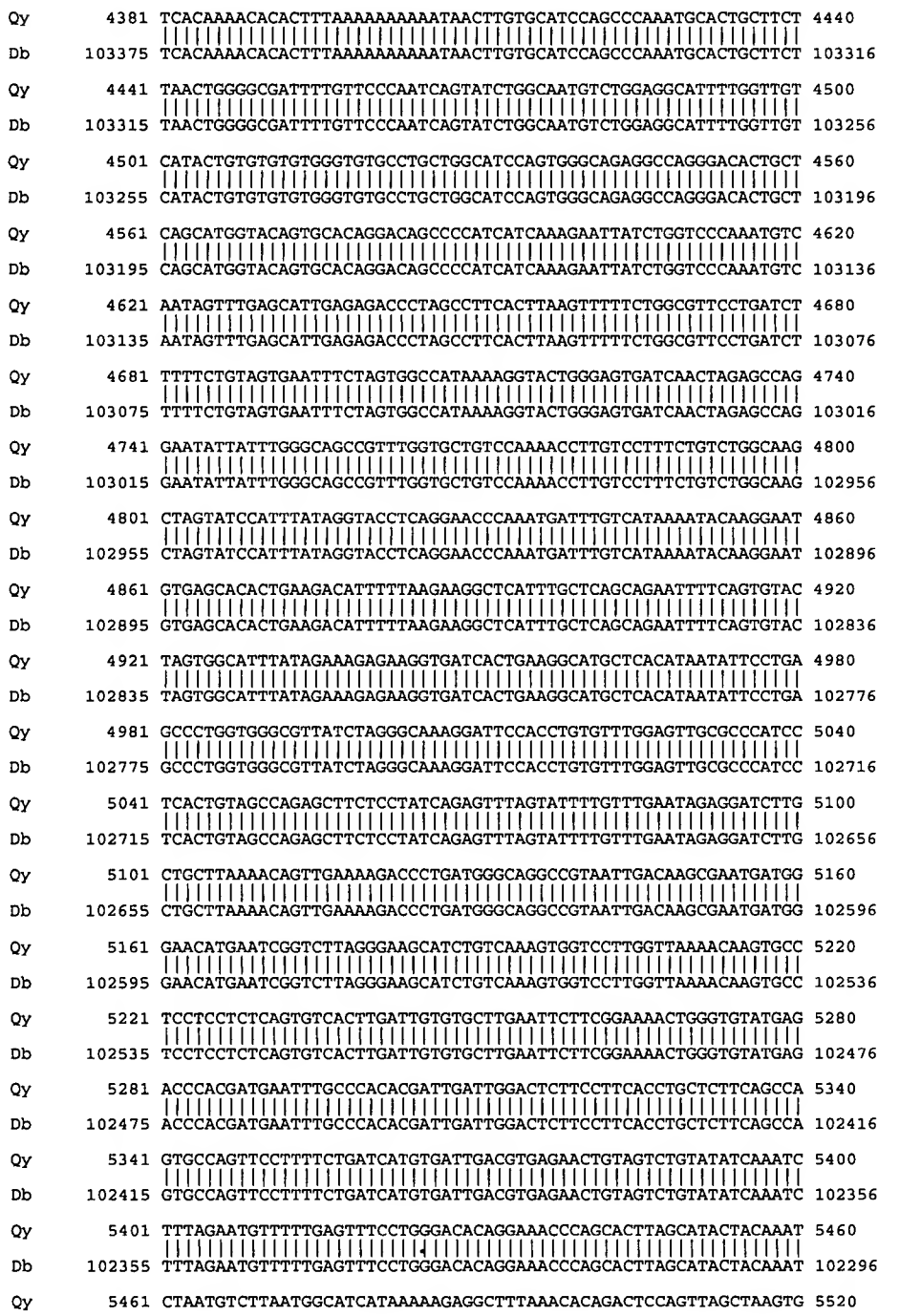
```
LOCUS       AF312915               121143 bp    DNA     linear   PRI 13-JUN-2001
DEFINITION  Homo sapiens chromosome 20 clones 97 and 127, complete sequence.
ACCESSION   AF312915
VERSION     AF312915.1  GI:11094030
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 121143)
  AUTHORS   Collins,C., Volik,S., Kowbel,D., Ginzinger,D., Ylstra,B.,
            Cloutier,T., Hawkins,T., Predki,P., Martin,C., Wernick,M.,
            Kuo,W.L., Alberts,A. and Gray,J.W.
  TITLE     Comprehensive genome sequence analysis of a breast cancer amplicon
  JOURNAL   Genome Res. 11 (6), 1034-1042 (2001)
  MEDLINE   21275464
   PUBMED   11381030
REFERENCE   2  (bases 1 to 121143)
  AUTHORS   Volik,S., Collins,C., Gray,J., Wernick,M., Kowbel,D., Stultz,K. and
            Martin,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-OCT-2000) Cancer Genetics, UCSF Cancer Center, 2340
            Sutter St., Rm. S151, San Francisco, CA 94706, USA
FEATURES             Location/Qualifiers
     source          1. .121143
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20q13.2"
                     /clone="97; 127; LBNL h117 and h79"
                     /note="isolated in Collins lab and sequenced by the Genome
                     Center at LBL"
     misc_feature    complement(1. .66093)
                     /note="overlaps Homo sapiens chromosome 20 clone
                     RP5-823G15 deposited in GenBank Accession Number
                     AL354993.19"
     repeat_region   194. .468
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   complement(513. .653)
                     /rpt_family="MER5"
                     /rpt_type=dispersed
     repeat_region   1053. .1153
                     /rpt_family="MER20"
                     /rpt_type=dispersed
     repeat_region   1254. .1330
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   complement(1356. .1676)
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   1681. .1866
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     misc_feature    2485. .5421
                     /note="similar to Homo sapiens chromosome 5 clone
                     CTC-269F1 deposited in GenBank Accession Number
                     AC011341.2"
```

```
repeat_region   2594. .2879
                /rpt_family="Alu"
                /rpt_type=dispersed
misc_feature    complement(2896. .4437)
                /note="similar to flow-sorted chromosome 20 HindIII
                fragment, SC20pF17C6 deposited in GenBank Accession Number
                Z94682.1"
repeat_region   3106. .3412
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   3889. .4206
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   4449. .4764
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(5671. .5987)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   6723. .7004
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(7486. .7625)
                /rpt_family="L1"
                /rpt_type=dispersed
repeat_region   7716. .7965
                /rpt_family="Alu"
                /rpt_type=dispersed
misc_feature    complement(8590. .118606)
                /note="similar to Mus musculus clone CT7-510C4 deposited
                in GenBank Accession Number AC023610.3"
misc_feature    complement(9852. .10109)
                /note="similar to STS WI-31010 deposited in GenBank
                Accession Number G24374.1"
mRNA            <10244. .>10904
                /product="hypothetical"
                /note="similar to Homo sapiens hypothetical protein
                OVC10-2 deposited in GenBank Accession Numbers AF230201.1,
                AP000365 and in NM_018692.1"
                /evidence=not_experimental
repeat_region   15301. .15588
                /rpt_family="Alu"
                /rpt_type=dispersed
misc_feature    16873. .17073
                /note="similar to Homo sapiens STS WI-16697 deposited in
                GenBank Accession Number G21490.1"
repeat_region   17577. .17857
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(19034. .19333)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   20136. .20719
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   20776. .21071
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(22344. .22638)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   22726. .22784
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(22984. .23268)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(23349. .23646)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(24008. .24189)
                /rpt_family="MIR"
                /rpt_type=dispersed
repeat_region   complement(24502. .24795)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   24883. .24941
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(25140. .25425)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(25507. .25805)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(26170. .26352)
                /rpt_family="MIR"
                /rpt_type=dispersed
repeat_region   26841. .27129
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   27270. .33290
                /rpt_family="L1"
                /rpt_type=dispersed
repeat_region   28512. .28730
                /rpt_family="MER25"
                /rpt_type=dispersed
repeat_region   34498. .35009
                /rpt_family="MLT2B2"
                /rpt_type=dispersed
repeat_region   35035. .35332
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   35343. .35522
                /rpt_family="Alu"
                /rpt_type=dispersed
misc_feature    35593. .121143
                /note="overlaps Homo sapiens clone RP4-724E16 on
                chromosome 20 deposited in GenBank Accession Number
                AL157838.24"
repeat_region   36137. .36250
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   36366. .36690
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   36949. .37020
                /rpt_family="MER5"
                /rpt_type=dispersed
repeat_region   complement(37692. .37766)
                /rpt_family="MER5"
                /rpt_type=dispersed
repeat_region   complement(38006. .38276)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   38863. .38984
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(39791. .40072)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(40100. .40395)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   40642. .41118
                /rpt_family="MER1"
                /rpt_type=dispersed
repeat_region   complement(41228. .41348)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   41649. .41938

Query Match           100.0%;  Score 10000;  DB 9;  Length 121143;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 10000;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```


Qy 1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 60
Db 107755 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 107696
Qy 61 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 120
Db 107695 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 107636
Qy 121 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 180
Db 107635 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 107576
Qy 181 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 240
Db 107575 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 107516
Qy 241 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300
Db 107515 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 107456
Qy 301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 360
Db 107455 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 107396
Qy 361 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 420
Db 107395 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 107336
Qy 421 AATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTTCAAAT 480
Db 107335 AATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTCAAAT 107276
Qy 481 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAAAACTCCAC 540
Db 107275 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAACCAAAAAACTCCAC 107216
Qy 541 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 600
Db 107215 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 107156
Qy 601 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 660
Db 107155 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 107096
Qy 661 TTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 720
Db 107095 TTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 107036
Qy 721 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 780
Db 107035 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 106976
Qy 781 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 840
Db 106975 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 106916
Qy 841 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 900
Db 106915 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 106856
Qy 901 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 960
Db 106855 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 106796
Qy 961 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 1020
Db 106795 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 106736
Qy 1021 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 1080
Db 106735 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 106676
Qy 1081 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 1140
Db 106675 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 106616
Qy 1141 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 1200
Db 106615 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 106556
Qy 1201 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 1260
Db 106555 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 106496
Qy 1261 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 1320
Db 106495 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 106436
Qy 1321 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 1380
Db 106435 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 106376
Qy 1381 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 1440
Db 106375 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 106316
Qy 1441 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 1500
Db 106315 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 106256
Qy 1501 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATGA 1560
Db 106255 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATGA 106196
Qy 1561 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 1620
Db 106195 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 106136
Qy 1621 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 1680
Db 106135 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 106076
Qy 1681 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 1740
Db 106075 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 106016
Qy 1741 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 1800
Db 106015 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 105956
Qy 1801 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
Db 105955 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 105896
Qy 1861 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 1920
Db 105895 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 105836
Qy 1921 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 1980
Db 105835 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 105776
Qy 1981 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 2040
Db 105775 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 105716
Qy 2041 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 2100
Db 105715 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 105656
Qy 2101 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 2160
Db 105655 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 105596
Qy 2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 2220

```
Db    105595 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 105536
Qy      2221 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 2280
Db    105535 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 105476
Qy      2281 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 2340
Db    105475 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 105416
Qy      2341 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTCT 2400
Db    105415 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTCT 105356
Qy      2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 2460
Db    105355 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 105296
Qy      2461 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520
Db    105295 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 105236
Qy      2521 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2580
Db    105235 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 105176
Qy      2581 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 2640
Db    105175 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 105116
Qy      2641 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2700
Db    105115 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 105056
Qy      2701 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2760
Db    105055 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 104996
Qy      2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2820
Db    104995 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 104936
Qy      2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2880
Db    104935 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 104876
Qy      2881 TTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 2940
Db    104875 TTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 104816
Qy      2941 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 3000
Db    104815 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 104756
Qy      3001 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 3060
Db    104755 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 104696
Qy      3061 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 3120
Db    104695 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 104636
Qy      3121 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 3180
Db    104635 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 104576
Qy      3181 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 3240
Db    104575 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 104516
Qy      3241 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 3300
Db    104515 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 104456
```

```
Qy      3301 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 3360
Db    104455 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 104396
Qy      3361 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 3420
Db    104395 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 104336
Qy      3421 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 3480
Db    104335 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 104276
Qy      3481 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 3540
Db    104275 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 104216
Qy      3541 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 3600
Db    104215 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 104156
Qy      3601 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 3660
Db    104155 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 104096
Qy      3661 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 3720
Db    104095 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 104036
Qy      3721 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 3780
Db    104035 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 103976
Qy      3781 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 3840
Db    103975 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 103916
Qy      3841 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 3900
Db    103915 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 103856
Qy      3901 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 3960
Db    103855 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 103796
Qy      3961 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 4020
Db    103795 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 103736
Qy      4021 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 4080
Db    103735 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 103676
Qy      4081 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 4140
Db    103675 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 103616
Qy      4141 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 4200
Db    103615 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 103556
Qy      4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 4260
Db    103555 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 103496
Qy      4261 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 4320
Db    103495 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 103436
Qy      4321 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 4380
Db    103435 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 103376
```


Qy 4381 TCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 4440
Db 103375 TCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 103316
Qy 4441 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 4500
Db 103315 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 103256
Qy 4501 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 4560
Db 103255 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 103196
Qy 4561 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 4620
Db 103195 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 103136
Qy 4621 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 4680
Db 103135 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 103076
Qy 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATCAACTAGAGCCAG 4740
Db 103075 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATCAACTAGAGCCAG 103016
Qy 4741 GAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGCAAG 4800
Db 103015 GAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGCAAG 102956
Qy 4801 CTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGGAAT 4860
Db 102955 CTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGGAAT 102896
Qy 4861 GTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTGTAC 4920
Db 102895 GTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTGTAC 102836
Qy 4921 TAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCCTGA 4980
Db 102835 TAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCCTGA 102776
Qy 4981 GCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCATCC 5040
Db 102775 GCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCATCC 102716
Qy 5041 TCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATCTTG 5100
Db 102715 TCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATCTTG 102656
Qy 5101 CTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGATGG 5160
Db 102655 CTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGATGG 102596
Qy 5161 GAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGTGCC 5220
Db 102595 GAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGTGCC 102536
Qy 5221 TCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTATGAG 5280
Db 102535 TCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTATGAG 102476
Qy 5281 ACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAGCCA 5340
Db 102475 ACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAGCCA 102416
Qy 5341 GTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAAATC 5400
Db 102415 GTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAAATC 102356
Qy 5401 TTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAAT 5460
Db 102355 TTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAAT 102296
Qy 5461 CTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAAGTG 5520
Db 102295 CTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAAGTG 102236
Qy 5521 GTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGAAAG 5580
Db 102235 GTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGAAAG 102176
Qy 5581 AAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGGCAG 5640
Db 102175 AAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGGCAG 102116
Qy 5641 GACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCTACC 5700
Db 102115 GACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCTACC 102056
Qy 5701 ACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCCCTG 5760
Db 102055 ACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCCCTG 101996
Qy 5761 TTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGCCTA 5820
Db 101995 TTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGCCTA 101936
Qy 5821 TTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCTTGA 5880
Db 101935 TTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCTTGA 101876
Qy 5881 AAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGCATT 5940
Db 101875 AAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGCATT 101816
Qy 5941 TCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGTCGG 6000
Db 101815 TCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGTCGG 101756
Qy 6001 AAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAAAAA 6060
Db 101755 AAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAAAAA 101696
Qy 6061 AGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAGTCT 6120
Db 101695 AGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAGTCT 101636
Qy 6121 AGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTAAAT 6180
Db 101635 AGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTAAAT 101576
Qy 6181 TTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTGGAT 6240
Db 101575 TTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTGGAT 101516
Qy 6241 AGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTTTGG 6300
Db 101515 AGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTTTGG 101456
Qy 6301 GGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTTGAA 6360
Db 101455 GGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTTGAA 101396
Qy 6361 GTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGTCCT 6420
Db 101395 GTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGTCCT 101336
Qy 6421 ATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTGTCAG 6480
Db 101335 ATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTGTCAG 101276
Qy 6481 ATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTGCTA 6540
Db 101275 ATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTGCTA 101216
Qy 6541 AATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAAGAT 6600

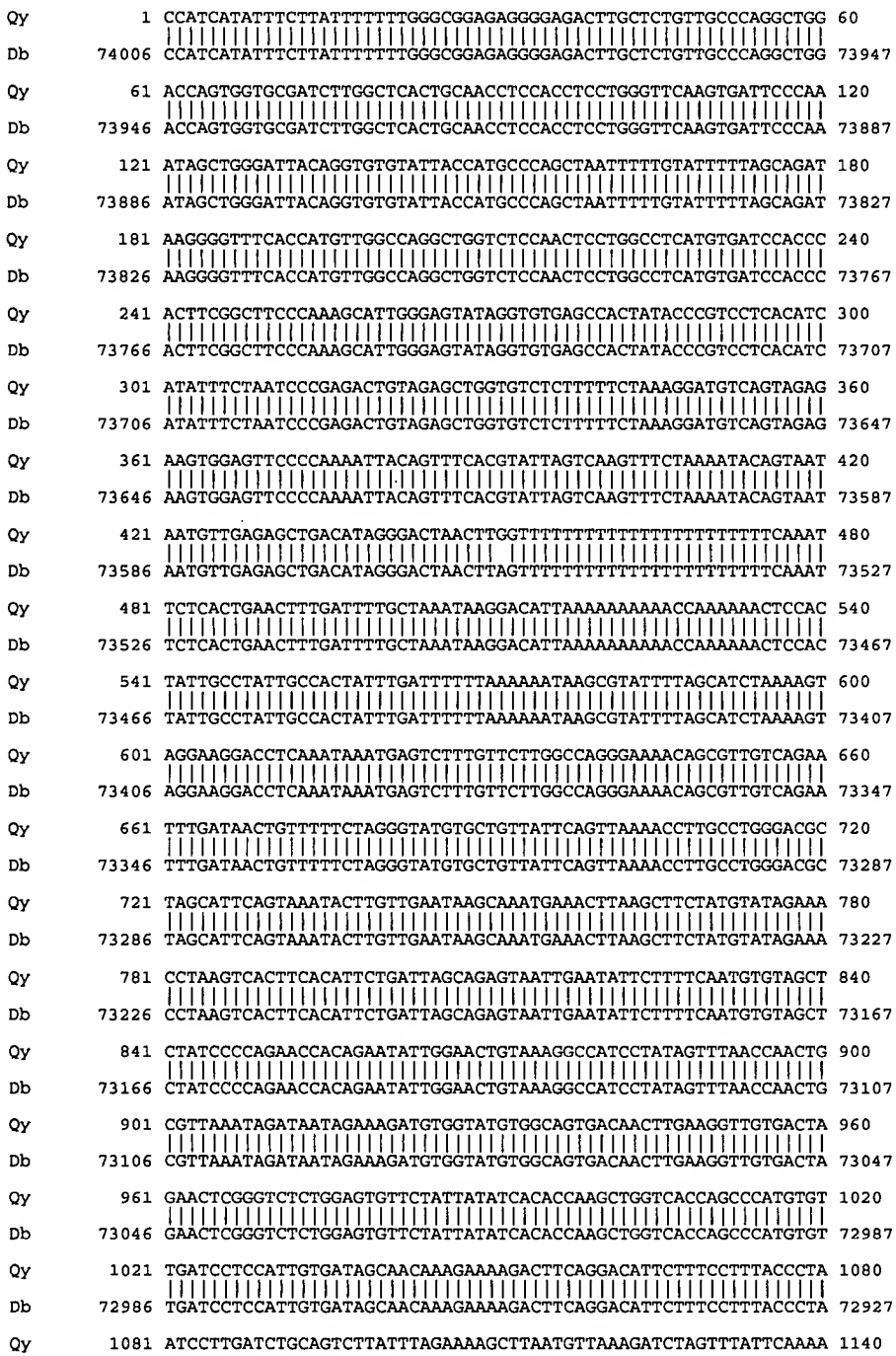
```
Db    101215 AATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAAGAT 101156

Qy      6601 TAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCTTAT 6660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101155 TAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCTTAT 101096

Qy      6661 TTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTATTC 6720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101095 TTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTATTC 101036

Qy      6721 TGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCTGCC 6780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101035 TGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCTGCC 100976

Qy      6781 TGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATCTTT 6840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100975 TGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATCTTT 100916

Qy      6841 CTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCAGAA 6900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100915 CTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCAGAA 100856

Qy      6901 AAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGCATT 6960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100855 AAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGCATT 100796

Qy      6961 TGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCATGCT 7020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100795 TGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCATGCT 100736

Qy      7021 TTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTCAGC 7080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100735 TTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTCAGC 100676

Qy      7081 TACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGAAAA 7140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100675 TACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGAAAA 100616

Qy      7141 ATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCAAAT 7200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100615 ATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCAAAT 100556

Qy      7201 TATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAACCA 7260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100555 TATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAACCA 100496

Qy      7261 TGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTCTGC 7320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100495 TGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTCTGC 100436

Qy      7321 CAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATTTTT 7380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100435 CAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATTTTT 100376

Qy      7381 ATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTTTCT 7440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100375 ATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTTTCT 100316

Qy      7441 TTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTGGTG 7500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100315 TTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTGGTG 100256

Qy      7501 CAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAGCCT 7560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100255 CAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAGCCT 100196

Qy      7561 CCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTTTAG 7620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100195 CCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTTTAG 100136

Qy      7621 TAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATC 7680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100135 TAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATC 100076

Qy      7681 CGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTGAGC 7740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100075 CGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTGAGC 100016

Qy      7741 CACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTTATC 7800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100015 CACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTTATC 99956

Qy      7801 TTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAATGAA 7860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99955 TTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAATGAA 99896

Qy      7861 ATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTAGGA 7920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99895 ATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTAGGA 99836

Qy      7921 GGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCAAGT 7980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99835 GGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCAAGT 99776

Qy      7981 TATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATATTTC 8040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99775 TATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATATTTC 99716

Qy      8041 CCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATTGAA 8100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99715 CCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATTGAA 99656

Qy      8101 GATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGAGAA 8160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99655 GATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGAGAA 99596

Qy      8161 ATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAATAA 8220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99595 ATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAATAA 99536

Qy      8221 ATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTTCTT 8280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99535 ATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTTCTT 99476

Qy      8281 GCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACA 8340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99475 GCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACA 99416

Qy      8341 TCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAA 8400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99415 TCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAA 99356

Qy      8401 GTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCG 8460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99355 GTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCG 99296

Qy      8461 CAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCT 8520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99295 CAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCT 99236

Qy      8521 GCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTC 8580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99235 GCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTC 99176

Qy      8581 TCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGAT 8640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99175 TCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGAT 99116

Qy      8641 AAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTT 8700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99115 AAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTT 99056

Qy      8701 GAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGAT 8760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99055 GAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGAT 98996
```

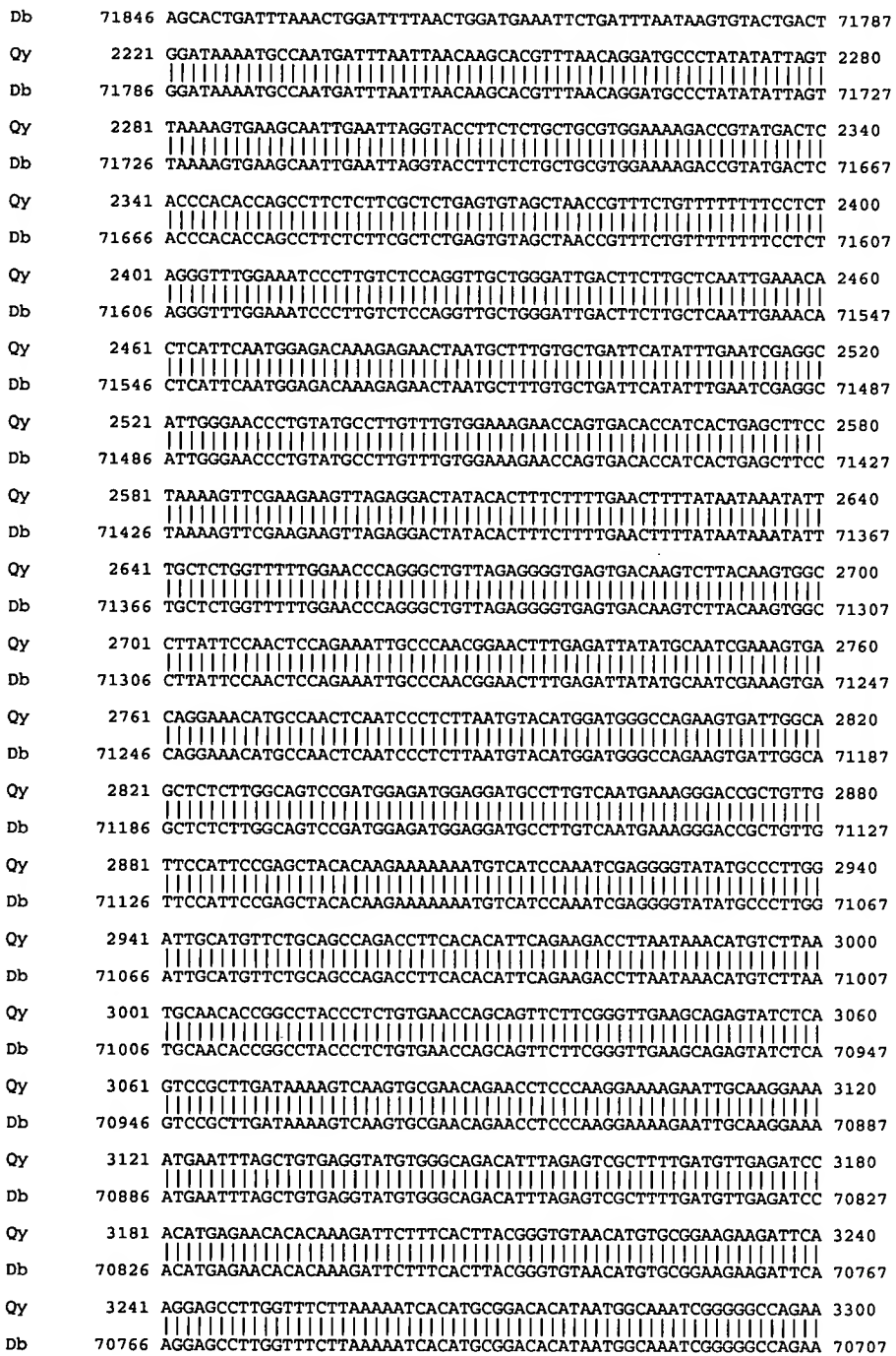
```
Qy       8761 GGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCT 8820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98995 GGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCT 98936

Qy       8821 GACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGG 8880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98935 GACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGG 98876

Qy       8881 GCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGT 8940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98875 GCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGT 98816

Qy       8941 CCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAG 9000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98815 CCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAG 98756

Qy       9001 CATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGA 9060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98755 CATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGA 98696

Qy       9061 CGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGT 9120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98695 CGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGT 98636

Qy       9121 AAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGG 9180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98635 AAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGG 98576

Qy       9181 AAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACT 9240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98575 AAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACT 98516

Qy       9241 TTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCC 9300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98515 TTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCC 98456

Qy       9301 GCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGAT 9360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98455 GCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGAT 98396

Qy       9361 AAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACC 9420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98395 AAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACC 98336

Qy       9421 CTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGG 9480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98335 CTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGG 98276

Qy       9481 GCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTT 9540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98275 GCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTT 98216

Qy       9541 GGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCC 9600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98215 GGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCC 98156

Qy       9601 GGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACA 9660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98155 GGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACA 98096

Qy       9661 TCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTC 9720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98095 TCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTC 98036

Qy       9721 CTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGC 9780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98035 CTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGC 97976

Qy       9781 TCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAA 9840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      97975 TCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAA 97916

Qy       9841 GGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAATCCA 9900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      97915 GGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAATCCA 97856

Qy       9901 GGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGTTCT 9960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      97855 GGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGTTCT 97796

Qy       9961 TGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000
             ||||||||||||||||||||||||||||||||||||||||
Db      97795 TGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 97756
```

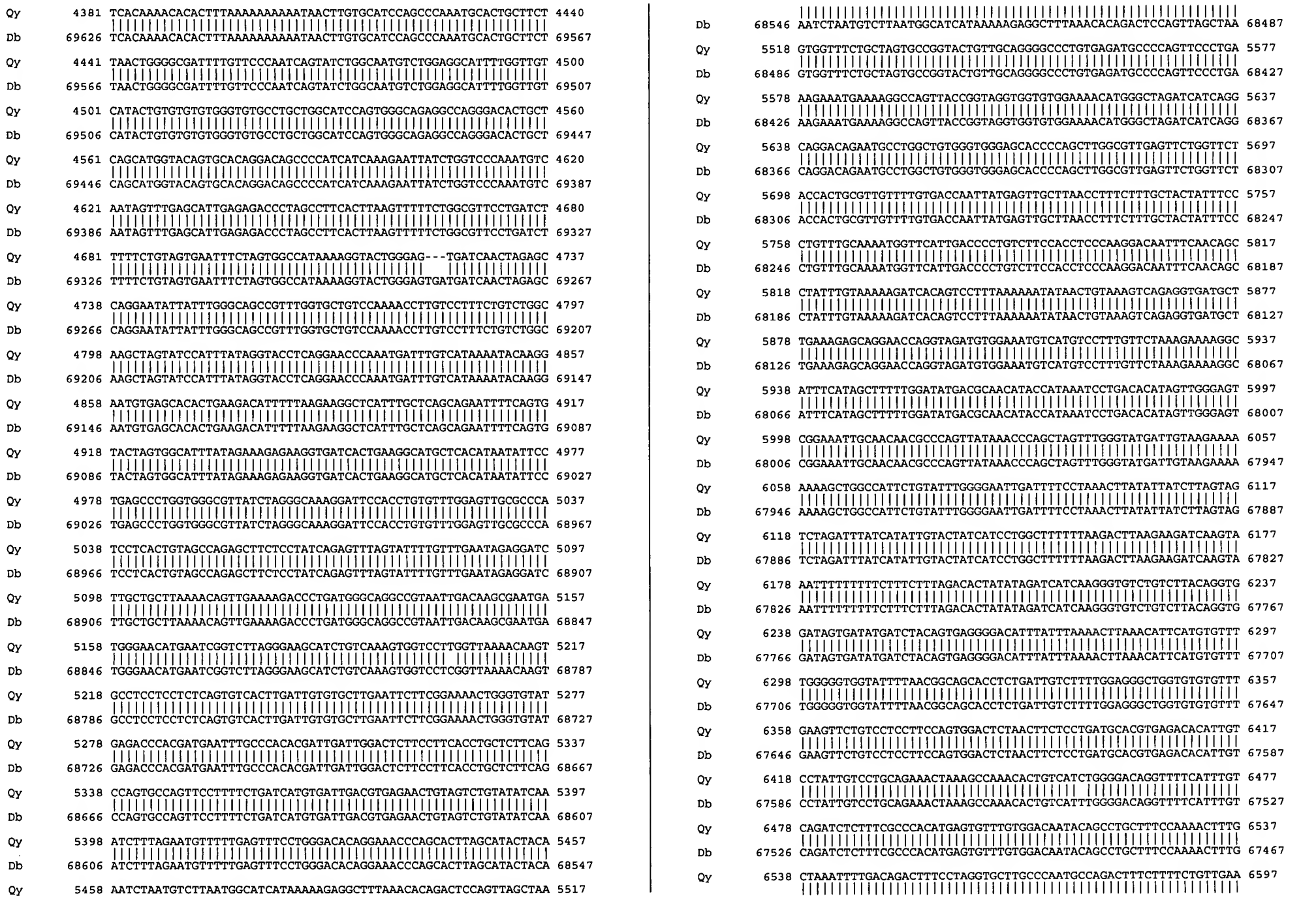
RESULT 3
AL157838/c

```
LOCUS       AL157838              128871 bp    DNA     linear   PRI 23-FEB-2001
DEFINITION  Human DNA sequence from clone RP4-724E16 on chromosome
            20q13.12-13.32 Contains the ZNF217 gene for zinc finger protein
            217, a novel gene, a putative novel gene, ESTs, GSSs, STSs and two
            CpG islands, complete sequence.
ACCESSION   AL157838
VERSION     AL157838.24  GI:9588158
KEYWORDS    HTG; CpG island; zinc finger; ZNF217.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 128871)
  AUTHORS   Wilson,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Jul 31, 2000 this sequence version replaced gi:9408255.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 20, constructed by the Sanger Centre Chromosome 20
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr20
            This sequence is the entire insert of clone RP4-724E16 The true
            right end of clone RP5-823G15 is at 19684 in this sequence. This
            sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. RP4-724E16 is from
            the library RPCI-4 constructed by the group of Pieter de Jong. For
            further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2.
FEATURES             Location/Qualifiers
     source          1. .128871
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="q13.12-13.32"
                     /clone="RP4-724E16"
                     /clone_lib="RPCI-4"
     repeat_region   1. .54
                     /note="MLT2A repeat: matches 362. .415 of consensus"
     repeat_region   55. .368
```

```
                     /note="AluSx repeat: matches 1. .312 of consensus"
repeat_region        369. .559
                     /note="AluSq repeat: matches 126. .312 of consensus"
repeat_region        1137. .1270
                     /note="FLAM_C repeat: matches 1. .133 of consensus"
misc_feature         complement(1396. .1840)
                     /note="match: GSS: Em:AQ261624"
repeat_region        1415. .1709
                     /note="AluJo repeat: matches 14. .288 of consensus"
repeat_region        1744. .1902
                     /note="MIR repeat: matches 14. .175 of consensus"
misc_feature         1871. .2326
                     /note="match: GSS: Em:AQ673688"
repeat_region        1948. .2041
                     /note="MER5A repeat: matches 6. .101 of consensus"
repeat_region        2622. .2787
                     /note="MER5A repeat: matches 4. .189 of consensus"
repeat_region        2806. .2870
                     /note="MER5A repeat: matches 10. .75 of consensus"
repeat_region        2872. .3026
                     /note="L1MA4 repeat: matches 6017. .6162 of consensus"
repeat_region        3027. .3307
                     /note="AluJo repeat: matches 1. .290 of consensus"
repeat_region        3308. .3436
                     /note="L1MA4 repeat: matches 6162. .6293 of consensus"
repeat_region        3455. .3551
                     /note="MER5A repeat: matches 61. .183 of consensus"
repeat_region        3564. .3697
                     /note="MER5B repeat: matches 14. .146 of consensus"
repeat_region        3871. .4004
                     /note="AluSq/x repeat: matches 1. .134 of consensus"
repeat_region        4249. .4274
                     /note="13 copies 2 mer gt 96% conserved"
repeat_region        4810. .5101
                     /note="AluSx repeat: matches 1. .297 of consensus"
repeat_region        5110. .5413
                     /note="AluSc repeat: matches 1. .305 of consensus"
repeat_region        5492. .5648
                     /note="L2 repeat: matches 2589. .2745 of consensus"
repeat_region        5659. .6135
                     /note="MER1A repeat: matches 1. .488 of consensus"
repeat_region        6164. .6225
                     /note="31 copies 2 mer aa 71% conserved"
repeat_region        6245. .6379
                     /note="AluSg/x repeat: matches 162. .296 of consensus"
repeat_region        6380. .6537
                     /note="79 copies 2 mer aa 60% conserved"
repeat_region        6658. .6967
                     /note="AluSg repeat: matches 1. .309 of consensus"
repeat_region        7320. .7542
                     /note="MER58B repeat: matches 117. .341 of consensus"
repeat_region        7545. .7838
                     /note="AluSx repeat: matches 1. .293 of consensus"
misc_feature         8465. .9068
                     /note="match: GSS: Em:AQ386102"
repeat_region        8523. .8814
                     /note="AluJb repeat: matches 1. .294 of consensus"
repeat_region        9262. .9561
                     /note="AluSq repeat: matches 1. .301 of consensus"
repeat_region        9562. .9599
                     /note="19 copies 2 mer tt 81% conserved"
repeat_region        9606. .9733
                     /note="MER33 repeat: matches 17. .148 of consensus"
repeat_region        9761. .9915
                     /note="MER5A repeat: matches 4. .183 of consensus"
repeat_region        9980. .10108
                     /note="L2 repeat: matches 2622. .2747 of consensus"
repeat_region        10128. .10437
                     /note="AluSx repeat: matches 2. .312 of consensus"
repeat_region        10654. .10724
                     /note="MIR repeat: matches 63. .137 of consensus"
repeat_region        10781. .11083
                     /note="AluSx repeat: matches 1. .306 of consensus"
misc_feature         complement(10897. .11319)
                     /note="match: GSS: Em:AQ499673"
misc_feature         complement(10991. .11319)
                     /note="match: GSS: Em:AQ283579"
repeat_region        11462. .11759
                     /note="AluSq repeat: matches 1. .299 of consensus"
repeat_region        11878. .12182
                     /note="AluSx repeat: matches 1. .300 of consensus"
repeat_region        12183. .12587
                     /note="L1PA8 repeat: matches 5757. .6163 of consensus"
repeat_region        13333. .13637
                     /note="MLT1G repeat: matches 3. .338 of consensus"
repeat_region        14201. .14507
                     /note="AluY repeat: matches 1. .308 of consensus"
repeat_region        14540. .14694
                     /note="L2 repeat: matches 2550. .2709 of consensus"
repeat_region        14764. .14868
                     /note="MIR repeat: matches 103. .208 of consensus"
repeat_region        14891. .15152
                     /note="AluJo repeat: matches 12. .299 of consensus"
repeat_region        15162. .15293
                     /note="L2 repeat: matches 1697. .1824 of consensus"
repeat_region        15317. .15544
                     /note="L1MB3 repeat: matches 5943. .6173 of consensus"
repeat_region        16306. .16603
                     /note="AluSc repeat: matches 1. .306 of consensus"
repeat_region        16610. .16901
                     /note="AluSx repeat: matches 1. .292 of consensus"
repeat_region        16989. .17035
                     /note="MLT1C repeat: matches 420. .463 of consensus"
repeat_region        17036. .17338
                     /note="AluJo repeat: matches 1. .301 of consensus"
repeat_region        17339. .17730
                     /note="MLT1C repeat: matches 1. .420 of consensus"
repeat_region        17958. .18269
                     /note="AluSx repeat: matches 1. .311 of consensus"
repeat_region        18288. .18315
                     /note="MER91 repeat: matches 32. .59 of consensus"
repeat_region        18383. .18507
                     /note="MIR repeat: matches 23. .148 of consensus"
repeat_region        18544. .18650
                     /note="MIR repeat: matches 137. .261 of consensus"
repeat_region        18799. .18820
                     /note="11 copies 2 mer aa 100% conserved"
repeat_region        18835. .19184
                     /note="THE1A repeat: matches 1. .354 of consensus"
misc_feature         19231. .19522
                     /note="match: STS: Em:HSA153YF5"
repeat_region        19297. .19346
                     /note="25 copies 2 mer ca 100% conserved"
repeat_region        19426. .19564
                     /note="MIR repeat: matches 6. .147 of consensus"
repeat_region        19624. .19916
                     /note="AluSx repeat: matches 1. .295 of consensus"
repeat_region        20005. .20204
                     /note="MER20 repeat: matches 1. .218 of consensus"
repeat_region        20215. .20317
                     /note="MER5A repeat: matches 4. .108 of consensus"
repeat_region        20410. .20541
                     /note="MER5B repeat: matches 1. .120 of consensus"
repeat_region        20544. .20687
                     /note="AluSg/x repeat: matches 154. .297 of consensus"
repeat_region        20689. .20814
                     /note="FLAM_A repeat: matches 14. .136 of consensus"
repeat_region        20816. .21218
                     /note="Charlie4a repeat: matches 17. .427 of consensus"
repeat_region        21337. .21634
                     /note="AluY repeat: matches 1. .297 of consensus"

Query Match             99.8%;  Score 9980.6;  DB 9;  Length 128871;
Best Local Similarity   99.9%;  Pred. No. 0;
Matches 9996;  Conservative    0;  Mismatches    4;  Indels    3;  Gaps    1;
```


Qy 1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 60
Db 74006 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 73947
Qy 61 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 120
Db 73946 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 73887
Qy 121 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 180
Db 73886 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 73827
Qy 181 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 240
Db 73826 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 73767
Qy 241 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300
Db 73766 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 73707
Qy 301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 360
Db 73706 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 73647
Qy 361 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 420
Db 73646 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 73587
Qy 421 AATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTTCAAAT 480
Db 73586 AATGTTGAGAGCTGACATAGGGACTAACTTAGTTTTTTTTTTTTTTTTTTTTTTTCAAAT 73527
Qy 481 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAAAACTCCAC 540
Db 73526 TCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAACCAAAAAACTCCAC 73467
Qy 541 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 600
Db 73466 TATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 73407
Qy 601 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 660
Db 73406 AGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGAA 73347
Qy 661 TTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 720
Db 73346 TTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACGC 73287
Qy 721 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 780
Db 73286 TAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAAA 73227
Qy 781 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 840
Db 73226 CCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGCT 73167
Qy 841 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 900
Db 73166 CTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACTG 73107
Qy 901 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 960
Db 73106 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 73047
Qy 961 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 1020
Db 73046 GAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTGT 72987
Qy 1021 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 1080
Db 72986 TGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTA 72927
Qy 1081 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 1140
Db 72926 ATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAA 72867
Qy 1141 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 1200
Db 72866 CTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCC 72807
Qy 1201 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 1260
Db 72806 TTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAA 72747
Qy 1261 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 1320
Db 72746 AGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTG 72687
Qy 1321 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 1380
Db 72686 GTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAA 72627
Qy 1381 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 1440
Db 72626 TGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTC 72567
Qy 1441 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 1500
Db 72566 CAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCAC 72507
Qy 1501 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATGA 1560
Db 72506 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATGA 72447
Qy 1561 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 1620
Db 72446 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCAC 72387
Qy 1621 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 1680
Db 72386 AAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGT 72327
Qy 1681 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 1740
Db 72326 GCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTT 72267
Qy 1741 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 1800
Db 72266 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTC 72207
Qy 1801 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
Db 72206 TGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 72147
Qy 1861 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 1920
Db 72146 ATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCT 72087
Qy 1921 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 1980
Db 72086 TGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTT 72027
Qy 1981 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 2040
Db 72026 AAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCT 71967
Qy 2041 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 2100
Db 71966 TTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGT 71907
Qy 2101 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 2160
Db 71906 ATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAA 71847
Qy 2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 2220


Db 71846 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACT 71787
Qy 2221 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 2280
Db 71786 GGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGT 71727
Qy 2281 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 2340
Db 71726 TAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACTC 71667
Qy 2341 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTCT 2400
Db 71666 ACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTCT 71607
Qy 2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 2460
Db 71606 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA 71547
Qy 2461 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520
Db 71546 CTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 71487
Qy 2521 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 2580
Db 71486 ATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCC 71427
Qy 2581 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 2640
Db 71426 TAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATT 71367
Qy 2641 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 2700
Db 71366 TGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGC 71307
Qy 2701 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 2760
Db 71306 CTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA 71247
Qy 2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 2820
Db 71246 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCA 71187
Qy 2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 2880
Db 71186 GCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTG 71127
Qy 2881 TTCCATTCCGAGCTACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 2940
Db 71126 TTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 71067
Qy 2941 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 3000
Db 71066 ATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAA 71007
Qy 3001 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 3060
Db 71006 TGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCA 70947
Qy 3061 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 3120
Db 70946 GTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAA 70887
Qy 3121 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 3180
Db 70886 ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCC 70827
Qy 3181 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 3240
Db 70826 ACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCA 70767
Qy 3241 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 3300
Db 70766 AGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAA 70707
Qy 3301 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 3360
Db 70706 GCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGC 70647
Qy 3361 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 3420
Db 70646 ACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTC 70587
Qy 3421 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 3480
Db 70586 CAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTA 70527
Qy 3481 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 3540
Db 70526 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCC 70467
Qy 3541 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 3600
Db 70466 TGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGAT 70407
Qy 3601 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 3660
Db 70406 GCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAA 70347
Qy 3661 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 3720
Db 70346 AAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACG 70287
Qy 3721 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 3780
Db 70286 ATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGC 70227
Qy 3781 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 3840
Db 70226 AAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCA 70167
Qy 3841 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 3900
Db 70166 AGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAA 70107
Qy 3901 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 3960
Db 70106 CCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGG 70047
Qy 3961 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 4020
Db 70046 AGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCG 69987
Qy 4021 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 4080
Db 69986 CCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTG 69927
Qy 4081 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 4140
Db 69926 AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTGT 69867
Qy 4141 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 4200
Db 69866 TCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACGC 69807
Qy 4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 4260
Db 69806 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTCT 69747
Qy 4261 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 4320
Db 69746 GTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTGA 69687
Qy 4321 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 4380
Db 69686 GTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTGC 69627


Qy 4381 TCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 4440
Db 69626 TCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTCT 69567
Qy 4441 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 4500
Db 69566 TAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTGT 69507
Qy 4501 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 4560
Db 69506 CATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGCT 69447
Qy 4561 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 4620
Db 69446 CAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGTC 69387
Qy 4621 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 4680
Db 69386 AATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATCT 69327
Qy 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAG---TGATCAACTAGAGC 4737
Db 69326 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATGATCAACTAGAGC 69267
Qy 4738 CAGGAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGC 4797
Db 69266 CAGGAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGGC 69207
Qy 4798 AAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGG 4857
Db 69206 AAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAGG 69147
Qy 4858 AATGTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTG 4917
Db 69146 AATGTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGTG 69087
Qy 4918 TACTAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCC 4977
Db 69086 TACTAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTCC 69027
Qy 4978 TGAGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCA 5037
Db 69026 TGAGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCCA 68967
Qy 5038 TCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATC 5097
Db 68966 TCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGATC 68907
Qy 5098 TTGCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGA 5157
Db 68906 TTGCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATGA 68847
Qy 5158 TGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAGT 5217
Db 68846 TGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTCGGTTAAAACAAGT 68787
Qy 5218 GCCTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTAT 5277
Db 68786 GCCTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTAT 68727
Qy 5278 GAGACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAG 5337
Db 68726 GAGACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCAG 68667
Qy 5338 CCAGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAA 5397
Db 68666 CCAGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAA 68607
Qy 5398 ATCTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACA 5457
Db 68606 ATCTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTACA 68547
Qy 5458 AATCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAA 5517
Db 68546 AATCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAA 68487
Qy 5518 GTGGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGA 5577
Db 68486 GTGGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGA 68427
Qy 5578 AAGAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGG 5637
Db 68426 AAGAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAGG 68367
Qy 5638 CAGGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCT 5697
Db 68366 CAGGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTCT 68307
Qy 5698 ACCACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCC 5757
Db 68306 ACCACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTCC 68247
Qy 5758 CTGTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGC 5817
Db 68246 CTGTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAGC 68187
Qy 5818 CTATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCT 5877
Db 68186 CTATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCT 68127
Qy 5878 TGAAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGC 5937
Db 68126 TGAAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGGC 68067
Qy 5938 ATTTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGT 5997
Db 68066 ATTTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAGT 68007
Qy 5998 CGGAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAA 6057
Db 68006 CGGAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAAA 67947
Qy 6058 AAAAGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAG 6117
Db 67946 AAAAGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTAG 67887
Qy 6118 TCTAGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTA 6177
Db 67886 TCTAGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGTA 67827
Qy 6178 AATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTG 6237
Db 67826 AATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGTG 67767
Qy 6238 GATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTT 6297
Db 67766 GATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTTT 67707
Qy 6298 TGGGGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTT 6357
Db 67706 TGGGGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTTT 67647
Qy 6358 GAAGTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGT 6417
Db 67646 GAAGTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTGT 67587
Qy 6418 CCTATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTGT 6477
Db 67586 CCTATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATTTGGGGACAGGTTTTCATTTGT 67527
Qy 6478 CAGATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTG 6537
Db 67526 CAGATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTTG 67467
Qy 6538 CTAAATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAA 6597

```
Db      67466 CTAAATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGAA 67407

Qy       6598 GATTAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCT 6657
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67406 GATTAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGCT 67347

Qy       6658 TATTTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTA 6717
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67346 TATTTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATTA 67287

Qy       6718 TTCTGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCT 6777
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67286 TTCTGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCCT 67227

Qy       6778 GCCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATC 6837
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67226 GCCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGATC 67167

Qy       6838 TTTCTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCA 6897
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67166 TTTCTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCCA 67107

Qy       6898 GAAAAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGC 6957
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67106 GAAAAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAGC 67047

Qy       6958 ATTTGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCAT 7017
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67046 ATTTGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCAT 66987

Qy       7018 GCTTTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTC 7077
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66986 GCTTTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTTC 66927

Qy       7078 AGCTACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGA 7137
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66926 AGCTACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGGA 66867

Qy       7138 AAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCA 7197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66866 AAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTCA 66807

Qy       7198 AATTATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAA 7257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66806 AATTATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTAA 66747

Qy       7258 CCATGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTC 7317
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66746 CCATGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACTC 66687

Qy       7318 TGCCAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATT 7377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66686 TGCCAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTATT 66627

Qy       7378 TTTATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTT 7437
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66626 TTTATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTTT 66567

Qy       7438 TCTTTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTG 7497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66566 TCTTTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGTG 66507

Qy       7498 GTGCAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAG 7557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66506 GTGCAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCAG 66447

Qy       7558 CCTCCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTT 7617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66446 CCTCCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTTT 66387

Qy       7618 TAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTG 7677
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66386 TAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTG 66327
```

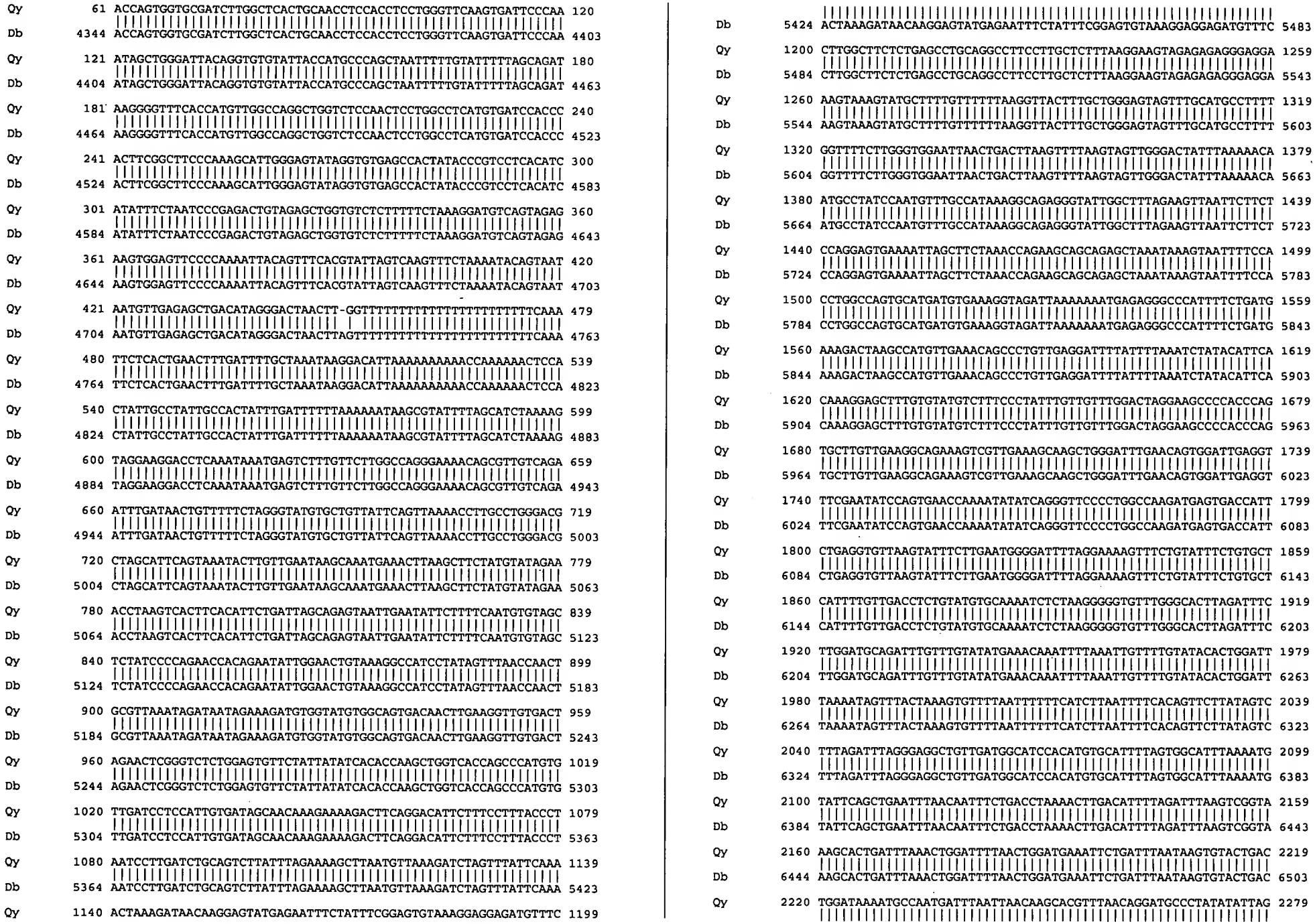
```
Qy       7678 ATCCGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTG 7737
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66326 ATCCGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGTG 66267

Qy       7738 AGCCACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTT 7797
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66266 AGCCACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCTT 66207

Qy       7798 ATCTTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAAT 7857
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66206 ATCTTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAAT 66147

Qy       7858 GAAATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTA 7917
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66146 GAAATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGTA 66087

Qy       7918 GGAGGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCA 7977
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66086 GGAGGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGCA 66027

Qy       7978 AGTTATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATAT 8037
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66026 AGTTATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATAT 65967

Qy       8038 TTCCCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATT 8097
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65966 TTCCCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTATT 65907

Qy       8098 GAAGATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGA 8157
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65906 GAAGATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATGA 65847

Qy       8158 GAAATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAA 8217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65846 GAAATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTAA 65787

Qy       8218 TAAATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTT 8277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65786 TAAATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATTT 65727

Qy       8278 CTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAG 8337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65726 CTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAG 65667

Qy       8338 ACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCT 8397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65666 ACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCT 65607

Qy       8398 GAAGTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGT 8457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65606 GAAGTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGT 65547

Qy       8458 GCGCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCA 8517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65546 GCGCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCA 65487

Qy       8518 CCTGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTC 8577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65486 CCTGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTC 65427

Qy       8578 CTCTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCT 8637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65426 CTCTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCT 65367

Qy       8638 GATAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCA 8697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65366 GATAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCA 65307

Qy       8698 GTTGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCG 8757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65306 GTTGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCG 65247
```

```
Qy       8758 GATGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCA 8817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65246 GATGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCA 65187

Qy       8818 GCTGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTG 8877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65186 GCTGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTG 65127

Qy       8878 GGGGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACC 8937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65126 GGGGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACC 65067

Qy       8938 TGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTG 8997
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65066 TGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTG 65007

Qy       8998 GAGCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGT 9057
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65006 GAGCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGT 64947

Qy       9058 CGACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTC 9117
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64946 CGACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCTCTAGTTTC 64887

Qy       9118 TGTAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAG 9177
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64886 TGTAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAG 64827

Qy       9178 GGGAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGC 9237
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64826 GGGAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGC 64767

Qy       9238 ACTTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGG 9297
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64766 ACTTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGG 64707

Qy       9298 GCCGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCG 9357
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64706 GCCGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCG 64647

Qy       9358 GATAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCC 9417
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64646 GATAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCC 64587

Qy       9418 ACCCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCG 9477
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64586 ACCCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCG 64527

Qy       9478 TGGGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAA 9537
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64526 TGGGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAA 64467

Qy       9538 TTTGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAG 9597
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64466 TTTGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAG 64407

Qy       9598 CCCGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATC 9657
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64406 CCCGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATC 64347

Qy       9658 ACATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGG 9717
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64346 ACATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGG 64287

Qy       9718 TTCCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGT 9777
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64286 TTCCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGT 64227

Qy       9778 GGCTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTA 9837
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64226 GGCTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTA 64167

Qy       9838 GAAGGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAAT 9897
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64166 GAAGGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAAT 64107

Qy       9898 CCAGGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGT 9957
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64106 CCAGGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAGT 64047

Qy       9958 TCTTGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000
             |||||||||||||||||||||||||||||||||||||||||||
Db      64046 TCTTGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 64004
```

RESULT 4
AC116668

```
LOCUS       AC116668              105023 bp    DNA     linear   HTG 01-MAY-2002
DEFINITION  Trypanosoma brucei chromosome V clone RPCI93-45E22, *** SEQUENCING
            IN PROGRESS ***, 3 unordered pieces.
ACCESSION   AC116668
VERSION     AC116668.5  GI:20376999
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Trypanosoma brucei
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 105023)
  AUTHORS   El-Sayed,N.M., Ghedin,E., Song,J., Larkin,C., Wanless,D., Jones,K.,
            Peterson,J., Hou,L., Zhao,H., Mason,T., Militscher,J., Pai,G., Van
            Aken,S., Utterback,T., Khalak,H.G., Gerard,C., Leech,V., Ullu,E.,
            Melville,S., White,O., Adams,M.D., Donelson,J.E. and Fraser,C.M.
  TITLE     Trypanosoma brucei GUTat10.1 RPCI93-45E22 BAC genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 105023)
  AUTHORS   El-Sayed,N.M., Khalak,H. and Adams,M.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2002) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
REFERENCE   3  (bases 1 to 105023)
  AUTHORS   El-Sayed,N.M., Khalak,H. and Adams,M.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2002) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
COMMENT     On May 1, 2002 this sequence version replaced gi:20340472.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    85781: contig of 85781 bp in length
            *    85782    85806: gap of unknown length
            *    85807   102956: contig of 17150 bp in length
            *   102957   102981: gap of unknown length
            *   102982   105023: contig of 2042 bp in length.
FEATURES             Location/Qualifiers
     source          1. .105023
                     /organism="Trypanosoma brucei"
                     /mol_type="genomic DNA"
                     /isolate="GUTat10.1"
                     /db_xref="taxon:5691"
                     /chromosome="V"
                     /clone="RPCI93-45E22"
ORIGIN
```

```
  Query Match             99.4%;  Score 9943.6;  DB 2;  Length 105023;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 9987;  Conservative    0;  Mismatches    9;  Indels    8;  Gaps    3;

Qy          1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4284 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGGAGACTTGCTCTGTTGCCCAGGCTGG 4343
```


Qy 61 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 120
Db 4344 ACCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCCCAA 4403
Qy 121 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 180
Db 4404 ATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAT 4463
Qy 181 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 240
Db 4464 AAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTGATCCACCC 4523
Qy 241 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300
Db 4524 ACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 4583
Qy 301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 360
Db 4584 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGTCAGTAGAG 4643
Qy 361 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 420
Db 4644 AAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAATACAGTAAT 4703
Qy 421 AATGTTGAGAGCTGACATAGGGACTAACTT-GGTTTTTTTTTTTTTTTTTTTTTTCAAA 479
Db 4704 AATGTTGAGAGCTGACATAGGGACTAACTTAGTTTTTTTTTTTTTTTTTTTTTTCAAA 4763
Qy 480 TTCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAAAACTCCA 539
Db 4764 TTCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAACCAAAAAACTCCA 4823
Qy 540 CTATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAG 599
Db 4824 CTATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAG 4883
Qy 600 TAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGA 659
Db 4884 TAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGTTGTCAGA 4943
Qy 660 ATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACG 719
Db 4944 ATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCCTGGGACG 5003
Qy 720 CTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAA 779
Db 5004 CTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATGTATAGAA 5063
Qy 780 ACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGC 839
Db 5064 ACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAATGTGTAGC 5123
Qy 840 TCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACT 899
Db 5124 TCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTAACCAACT 5183
Qy 900 GCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACT 959
Db 5184 GCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACT 5243
Qy 960 AGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTG 1019
Db 5244 AGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGCCCATGTG 5303
Qy 1020 TTGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCT 1079
Db 5304 TTGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCT 5363
Qy 1080 AATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAA 1139
Db 5364 AATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAA 5423
Qy 1140 ACTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTC 1199
Db 5424 ACTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTC 5483
Qy 1200 CTTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGA 1259
Db 5484 CTTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGA 5543
Qy 1260 AAGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTT 1319
Db 5544 AAGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTT 5603
Qy 1320 GGTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACA 1379
Db 5604 GGTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACA 5663
Qy 1380 ATGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCT 1439
Db 5664 ATGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCT 5723
Qy 1440 CCAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCA 1499
Db 5724 CCAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCA 5783
Qy 1500 CCTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATG 1559
Db 5784 CCTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATG 5843
Qy 1560 AAAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCA 1619
Db 5844 AAAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCA 5903
Qy 1620 CAAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAG 1679
Db 5904 CAAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAG 5963
Qy 1680 TGCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGT 1739
Db 5964 TGCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGT 6023
Qy 1740 TTCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATT 1799
Db 6024 TTCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATT 6083
Qy 1800 CTGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCT 1859
Db 6084 CTGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCT 6143
Qy 1860 CATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTC 1919
Db 6144 CATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTC 6203
Qy 1920 TTGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATT 1979
Db 6204 TTGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATT 6263
Qy 1980 TAAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTC 2039
Db 6264 TAAAATAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTC 6323
Qy 2040 TTTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATG 2099
Db 6324 TTTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATG 6383
Qy 2100 TATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTA 2159
Db 6384 TATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTA 6443
Qy 2160 AAGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGAC 2219
Db 6444 AAGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGAC 6503
Qy 2220 TGGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAG 2279

```
Db       6504 TGGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAG 6563

Qy       2280 TTAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACT 2339

Db       6564 TTAAAAGTGAAGCAATTGAATTAGGTACCTTCTCTGCTGCGTGGAAAAGACCGTATGACT 6623

Qy       2340 CACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTC 2399

Db       6624 CACCCACACCAGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTTTCTGTTTTTTTTCCTC 6683

Qy       2400 TAGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAAC 2459

Db       6684 TAGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAAC 6743

Qy       2460 ACTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGG 2519

Db       6744 ACTCATTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGG 6803

Qy       2520 CATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTC 2579

Db       6804 CATTGGGAACCCTGTATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTC 6863

Qy       2580 CTAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATAT 2639

Db       6864 CTAAAAGTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATAT 6923

Qy       2640 TTGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGG 2699

Db       6924 TTGCTCTGGTTTTTGGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGG 6983

Qy       2700 CCTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTG 2759

Db       6984 CCTTATTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTG 7043

Qy       2760 ACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGC 2819

Db       7044 ACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGC 7103

Qy       2820 AGCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTT 2879

Db       7104 AGCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTT 7163

Qy       2880 GTTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTG 2939

Db       7164 GTTCCATTCCGAGCTACACAAGAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTG 7223

Qy       2940 GATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTA 2999

Db       7224 GATTGCATGTTCTGCAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTA 7283

Qy       3000 ATGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTC 3059

Db       7284 ATGCAACACCGGCCTACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTC 7343

Qy       3060 AGTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAA 3119

Db       7344 AGTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAA 7403

Qy       3120 AATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATC 3179

Db       7404 AATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATC 7463

Qy       3180 CACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTC 3239

Db       7464 CACATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTC 7523

Qy       3240 AAGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGA 3299

Db       7524 AAGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGA 7583

Qy       3300 AGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTG 3359

Db       7584 AGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTG 7643

Qy       3360 CACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTT 3419

Db       7644 CACGCGGCCGAGAGCATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTT 7703

Qy       3420 CCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGT 3479

Db       7704 CCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGT 7763

Qy       3480 ACCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTC 3539

Db       7764 ACCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTC 7823

Qy       3540 CTGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGA 3599

Db       7824 CTGCAGTTGTTCAACTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGA 7883

Qy       3600 TGCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGA 3659

Db       7884 TGCATCCCTCAGCTCGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGA 7943

Qy       3660 AAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGAC 3719

Db       7944 AAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGAC 8003

Qy       3720 GATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCG 3779

Db       8004 GATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCG 8063

Qy       3780 CAAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCC 3839

Db       8064 CAAGAGAAAGAGAAGTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCC 8123

Qy       3840 AAGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGA 3899

Db       8124 AAGTTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGA 8183

Qy       3900 ACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCG 3959

Db       8184 ACCTACCACCAGCTGGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCG 8243

Qy       3960 GAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCC 4019

Db       8244 GAGTCGCCCACCATGTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCC 8303

Qy       4020 GCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCT 4079

Db       8304 GCCCCTCTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCT 8363

Qy       4080 GAGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTG 4139

Db       8364 GAGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCCGTGCTG 8423

Qy       4140 TTCCGCCTGTGTCTGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACG 4199

Db       8424 TTCCGCCTG----TGTCTGTCTCCCCGTCTCCCCCTCTCTATTCCCATCTCCAGACAACG 8479

Qy       4200 CTGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTC 4259

Db       8480 CTGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTTGGTATCACTC 8539

Qy       4260 TGTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTG 4319

Db       8540 TGTGTAAGTCATTTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAATAACAGGGTTG 8599

Qy       4320 AGTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTG 4379

Db       8600 AGTTAAGAGGTCTCCTTGTTCTGAAAATATATATATATTTTTTAAACGTGTATCGTTTTG 8659

Qy       4380 CTCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTC 4439

Db       8660 CTCACAAAACACACTTTAAAAAAAAAATAACTTGTGCATCCAGCCCAAATGCACTGCTTC 8719
```

```
Qy      4440 TTAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTG 4499
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8720 TTAACTGGGGCGATTTTGTTCCCAATCAGTATCTGGCAATGTCTGGAGGCATTTTGGTTG 8779

Qy      4500 TCATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGC 4559
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8780 TCATACTGTGTGTGTGGGTGTGCCTGCTGGCATCCAGTGGGCAGAGGCCAGGGACACTGC 8839

Qy      4560 TCAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGT 4619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8840 TCAGCATGGTACAGTGCACAGGACAGCCCCATCATCAAAGAATTATCTGGTCCCAAATGT 8899

Qy      4620 CAATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATC 4679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8900 CAATAGTTTGAGCATTGAGAGACCCTAGCCTTCACTTAAGTTTTTCTGGCGTTCCTGATC 8959

Qy      4680 TTTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAG---TGATCAACTAGAG 4736
             |||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db      8960 TTTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGGAGTGATGATCAACTAGAG 9019

Qy      4737 CCAGGAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGG 4796
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9020 CCAGTAATATTATTTGGGCAGCCGTTTGGTGCTGTCCAAAACCTTGTCCTTTCTGTCTGG 9079

Qy      4797 CAAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAG 4856
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9080 CAAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCATAAAATACAAG 9139

Qy      4857 GAATGTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGT 4916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9140 GAATGTGAGCACACTGAAGACATTTTTAAGAAGGCTCATTTGCTCAGCAGAATTTTCAGT 9199

Qy      4917 GTACTAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTC 4976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9200 GTACTAGTGGCATTTATAGAAAGAGAAGGTGATCACTGAAGGCATGCTCACATAATATTC 9259

Qy      4977 CTGAGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCC 5036
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9260 CTGAGCCCTGGTGGGCGTTATCTAGGGCAAAGGATTCCACCTGTGTTTGGAGTTGCGCCC 9319

Qy      5037 ATCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGAT 5096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9320 ATCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTTTGAATAGAGGAT 9379

Qy      5097 CTTGCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATG 5156
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9380 CTTGCTGCTTAAAACAGTTGAAAAGACCCTGATGGGCAGGCCGTAATTGACAAGCGAATG 9439

Qy      5157 ATGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTTGGTTAAAACAAG 5216
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9440 ATGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCCTCGGTTAAAACAAG 9499

Qy      5217 TGCCTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTA 5276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9500 TGCCTCCTCCTCTCAGTGTCACTTGATTGTGTGCTTGAATTCTTCGGAAAACTGGGTGTA 9559

Qy      5277 TGAGACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCA 5336
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9560 TGAGACCCACGATGAATTTGCCCACACGATTGATTGGACTCTTCCTTCACCTGCTCTTCA 9619

Qy      5337 GCCAGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCA 5396
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9620 GCCAGTGCCAGTTCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCA 9679

Qy      5397 AATCTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTAC 5456
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9680 AATCTTTAGAATGTTTTTGAGTTTCCTGGGACACAGGAAACCCAGCACTTAGCATACTAC 9739

Qy      5457 AAATCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTA 5516
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9740 AAATCTAATGTCTTAATGGCATCATAAAAAGAGGCTTTAAACACAGACTCCAGTTAGCTA 9799

Qy      5517 AGTGGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTG 5576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9800 AGTGGTTTCTGCTAGTGCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTG 9859

Qy      5577 AAAGAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAG 5636
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9860 AAAGAAATGAAAAGGCCAGTTACCGGTAGGTGGTGTGGAAAACATGGGCTAGATCATCAG 9919

Qy      5637 GCAGGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTC 5696
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9920 GCAGGACAGAATGCCTGGCTGTGGGTGGGAGCACCCCAGCTTGGCGTTGAGTTCTGGTTC 9979

Qy      5697 TACCACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTC 5756
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9980 TACCACTGCGTTGTTTTGTGACCAATTATGAGTTGCTTAACCTTTCTTTGCTACTATTTC 10039

Qy      5757 CCTGTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAG 5816
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10040 CCTGTTTGCAAAATGGTTCATTGACCCCTGTCTTCCACCTCCCAAGGACAATTTCAACAG 10099

Qy      5817 CCTATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGC 5876
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10100 CCTATTTGTAAAAAGATCACAGTCCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGC 10159

Qy      5877 TTGAAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGG 5936
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10160 TTGAAAGAGCAGGAACCAGGTAGATGTGGAAATGTCATGTCCTTTGTTCTAAAGAAAAGG 10219

Qy      5937 CATTTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAG 5996
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10220 CATTTCATAGCTTTTTGGATATGACGCAACATACCATAAATCCTGACACATAGTTGGGAG 10279

Qy      5997 TCGGAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAA 6056
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10280 TCGGAAATTGCAACAACGCCCAGTTATAAACCCAGCTAGTTTGGGTATGATTGTAAGAAA 10339

Qy      6057 AAAAAGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTA 6116
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10340 AAAAAGCTGGCCATTCTGTATTTGGGGAATTGATTTTCCTAAACTTATATTATCTTAGTA 10399

Qy      6117 GTCTAGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGT 6176
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10400 GTCTAGATTTATCATATTGTACTATCATCCTGGCTTTTTTAAGACTTAAGAAGATCAAGT 10459

Qy      6177 AAATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGT 6236
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10460 AAATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTTACAGGT 10519

Qy      6237 GGATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTT 6296
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10520 GGATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAACATTCATGTGTT 10579

Qy      6297 TTGGGGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTT 6356
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10580 TTGGGGGTGGTATTTTAACGGCAGCACCTCTGATTGTCTTTTGGAGGGCTGGTGTGTGTT 10639

Qy      6357 TGAAGTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTG 6416
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10640 TGAAGTTCTGTCCTCCTTCCAGTGGACTCTAACTTCTCCTGATGCACGTGAGACACATTG 10699

Qy      6417 TCCTATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATCTGGGGACAGGTTTTCATTTG 6476
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10700 TCCTATTGTCCTGCAGAAACTAAAGCCAAACACTGTCATTTGGGGACAGGTTTTCATTTG 10759

Qy      6477 TCAGATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTT 6536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10760 TCAGATCTCTTTCGCCCACATGAGTGTTTGTGGACAATACAGCCTGCTTTCCAAAACTTT 10819

Qy      6537 GCTAAATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGA 6596
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10820 GCTAAATTTTGACAGACTTTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTTGA 10879

Qy      6597 AGATTAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGC 6656
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      10880 AGATTAAGTTGTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAATCCTAACCTTAAACGGC 10939

Qy       6657 TTATTTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATT 6716

Db      10940 TTATTTTTCCCCTGGTGGTTGGGAAGTTGACGGTTTGTAATTGGCTCATTTTTCTAAATT 10999

Qy       6717 ATTCTGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCC 6776

Db      11000 ATTCTGAAGAAGATAATTTTTCCCGCCAGTATGTATGTCCACCTTCAGTTTGCCAGATCC 11059

Qy       6777 TGCCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGAT 6836

Db      11060 TGCCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGGCAATTCAGTCTATGAAATGAT 11119

Qy       6837 CTTTCTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCC 6896

Db      11120 CTTTCTTGTGATTAAGGCAAACGAAGAACTGAATGTTTAATAGTGTACTCTGCTGTACCC 11179

Qy       6897 AGAAAAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAG 6956

Db      11180 AGAAAAAAACAAAACAAAATCATGTTATAACACTCTAAAACTTCAAACAACCTCCAACAG 11239

Qy       6957 CATTTGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCA 7016

Db      11240 CATTTGGTGTGTGTCTAGCCGTTTTGTTCTAACCCGATGTTATATAAAAGAATTTTTTCA 11299

Qy       7017 TGCTTTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTT 7076

Db      11300 TGCTTTCCAAAAATGTTTATGTCAAGAATATTTAAGTCAGCATGCCTTATTCAGGTACTT 11359

Qy       7077 CAGCTACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGG 7136

Db      11360 CAGCTACCTTCTTATATAAATATTTTTGTTTTTCCTTTAAGATAAAAATGATGATGGAGG 11419

Qy       7137 AAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTC 7196

Db      11420 AAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTCCGTTC 11479

Qy       7197 AAATTATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTA 7256

Db      11480 AAATTATTACCTCAATATTCATCTCAGAACGCATACAGGTAAAGAACTTTTATTTTTTTA 11539

Qy       7257 ACCATGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACT 7316

Db      11540 ACCATGCATTAGTTAAATTATGTAGTTATCTAATTTTTTTGTTGTTGTTGTTCAGATACT 11599

Qy       7317 CTGCCAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTAT 7376

Db      11600 CTGCCAGATCCTTGGACTAGCTTAAGGATAAATATGTAGCATGTTGATTGCAGTGGTTAT 11659

Qy       7377 TTTTATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTT 7436

Db      11660 TTTTATTCTTTTAGTGCCATTGTAACTTGAGCCATTGTTCTTATTTGCAGTTCATTTCTT 11719

Qy       7437 TTCTTTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGT 7496

Db      11720 TTCTTTCTTTTTTGTTTTTTGAGACGGAGTCTTGCTCTGTCACCTCGGCTGGAGTGCAGT 11779

Qy       7497 GGTGCAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCA 7556

Db      11780 GGTGCAATTTCGGCTCACTGCAGCCTCCACCTCCCTGGTTCAAGCAATACTCCTGCCTCA 11839

Qy       7557 GCCTCCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTT 7616

Db      11840 GCCTCCCCAGTAGTTGGGATTACAGGTACCTGCCACCACACCCGGCTAATTTCTGTATTT 11899

Qy       7617 TTAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGT 7676

Db      11900 TTAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGT 11959

Qy       7677 GATCCGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGT 7736

Db      11960 GATCCGCTCACCTTGGCCTCCCATAGTGTTGGCCTCCCATAGTGCTGGGATTACAGGCGT 12019
```

```
Qy       7737 GAGCCACCGCGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCT 7796

Db      12020 GAGCCACCGTGCCCGGACAAAGTTCATTTGTTTAGTTTATGACTGCTATGTCCTGACTCT 12079

Qy       7797 TATCTTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAA 7856

Db      12080 TATCTTATTAAAAGCTACAGTATTTTAAAATGCTGCATCTTATGTCTTTATGATTGAGAA 12139

Qy       7857 TGAAATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGT 7916

Db      12140 TGAAATGAGAATCTATTTAGTAGTCTTGAGATTGTGAAAGGAGCTATGACATCATGATGT 12199

Qy       7917 AGGAGGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGC 7976

Db      12200 AGGAGGCTGCGTAGATTTGAAATTTCATCTCTTCCACTTACTATCTGTGCACCCTTGGGC 12259

Qy       7977 AAGTTATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATA 8036

Db      12260 AAGTTATTTAACCTTTTTGTGCTTTTAGTTTTCTTTGCTGTAAAAGTAGAATAATACATA 12319

Qy       8037 TTTCCCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTAT 8096

Db      12320 TTTCCCTAGGGCTGTTAGGAAGATTAAATAAGTTAGAAGTGTTGCTGTTAATTTTTCTAT 12379

Qy       8097 TGAAGATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATG 8156

Db      12380 TGAAGATAGGCATTCATAATTTCAAATATTCATTACAGTAAGGATGATAAAGAACTGATG 12439

Qy       8157 AGAAATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTA 8216

Db      12440 AGAAATCCTATGTGATAGTAGATCGAGAAAGCAAAAGGAGGAAAGAAGCCTGTTTTCTTA 12499

Qy       8217 ATAAATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATT 8276

Db      12500 ATAAATAGATATTTGATCTATTTCAGTGCTTTTCATACACTTCTATAATAAAGTGCCATT 12559

Qy       8277 TCTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAA 8336

Db      12560 TCTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAA 12619

Qy       8337 GACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGC 8396

Db      12620 GACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGC 12679

Qy       8397 TGAAGTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAG 8456

Db      12680 TGAAGTCAAGAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAG 12739

Qy       8457 TGCGCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCC 8516

Db      12740 TGCGCAAACCAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCC 12799

Qy       8517 ACCTGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGT 8576

Db      12800 ACCTGCAAAGCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGT 12859

Qy       8577 CCTCTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGC 8636

Db      12860 CCTCTCACCAGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGC 12919

Qy       8637 TGATAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGC 8696

Db      12920 TGATAAAGTGAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGC 12979

Qy       8697 AGTTGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCC 8756

Db      12980 AGTTGAAACTCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCC 13039

Qy       8757 GGATGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGC 8816

Db      13040 GGATGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGC 13099
```

```
Qy       8817 AGCTGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGT 8876
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13100 AGCTGACTGCAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGT 13159

Qy       8877 GGGGGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCAC 8936
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13160 GGGGGCTCTTCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCAC 13219

Qy       8937 CTGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACT 8996
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13220 CTGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACT 13279

Qy       8997 GGAGCATAAATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAG 9056
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      13280 GGAGCATAAATACAATCCTGACATTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAG 13339

Qy       9057 TCGACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTT 9116
             ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db      13340 TCGACGTACCGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCTCTAGTTT 13399

Qy       9117 CTGTAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAA 9176
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13400 CTGTAAACCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAA 13459

Qy       9177 GGGGAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAG 9236
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13460 GGGGAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAG 13519

Qy       9237 CACTTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGG 9296
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13520 CACTTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGG 13579

Qy       9297 GGCCGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACC 9356
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13580 GGCCGCCACCAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACC 13639

Qy       9357 GGATAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCC 9416
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13640 GGATAAGACAAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCC 13699

Qy       9417 CACCCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCC 9476
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db      13700 CACCCTCGGCAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGGCAGCCC 13759

Qy       9477 GTGGGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGA 9536
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      13760 GTGGGCACCTCCGGGAAGAGACTATTTCTGTAATCAGAGTGCCAGCAATACTGCAGCAGA 13819

Qy       9537 ATTTGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCA 9596
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13820 ATTTGGTGAGCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCA 13879

Qy       9597 GCCCGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCAT 9656
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13880 GCCCGGGGCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCAT 13939

Qy       9657 CACATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAG 9716
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13940 CACATCACTGTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAG 13999

Qy       9717 GTTCCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGG 9776
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14000 GTTCCTGAGCTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGG 14059

Qy       9777 TGGCTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTT 9836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14060 TGGCTCCGGGCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTT 14119

Qy       9837 AGAAGGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAA 9896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14120 AGAAGGTATTGCATGAGGGGCGTCGTGTTTAAATGGCTGCCTACAGTGATTAATAGCTAA 14179

Qy       9897 TCCAGGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAG 9956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14180 TCCAGGCATTCTCAGTGGAGATGGTACCACTCCCAAGGGTGGGGGGTAGGCAGCCAGAAG 14239

Qy       9957 TTCTTGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 10000
             |||||||||||||||||||||||||||||||||||||||||||||
Db      14240 TTCTTGGGGGTCACAGAGAGAAGCATTCTTAGATACGGCAGTGG 14283
```

RESULT 5
AC015713/c
LOCUS AC015713 165471 bp DNA linear HTG 24-AUG-2002
DEFINITION Homo sapiens chromosome 22 clone RP11-67H19 map 22, WORKING DRAFT SEQUENCE, 24 unordered pieces.
ACCESSION AC015713
VERSION AC015713.3 GI:7329268
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 165471)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 22, clone RP11-67H19
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 165471)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 165471)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Mar 26, 2000 this sequence version replaced gi:6514042.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
---------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

```
    Center code: WIBR
    Web site: http://www-seq.wi.mit.edu
    Contact: sequence_submissions@genome.wi.mit.edu
-------------- Project Information
    Center project name: L1398
    Center clone name: 67_H_19
-------------- Summary Statistics
    Sequencing vector: M13; M77815; 100% of reads
    Chemistry: Dye-terminator Big Dye; 100% of reads
    Assembly program: Phrap; version 0.960731
    Consensus quality: 155715 bases at least Q40
    Consensus quality: 160515 bases at least Q30
    Consensus quality: 162183 bases at least Q20
    Insert size: 138000; agarose-fp
    Insert size: 163171; sum-of-contigs
    Quality coverage: 6.8 in Q20 bases; agarose-fp
    Quality coverage: 5.8 in Q20 bases; sum-of-contigs
--------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1323: contig of 1323 bp in length
*     1324     1423: gap of 100 bp
*     1424     2727: contig of 1304 bp in length
*     2728     2827: gap of 100 bp
*     2828     4344: contig of 1517 bp in length
*     4345     4444: gap of 100 bp
*     4445     5722: contig of 1278 bp in length
*     5723     5822: gap of 100 bp
*     5823     6884: contig of 1062 bp in length
*     6885     6984: gap of 100 bp
*     6985     8290: contig of 1306 bp in length
*     8291     8390: gap of 100 bp
*     8391     9932: contig of 1542 bp in length
*     9933    10032: gap of 100 bp
*    10033    12359: contig of 2327 bp in length
*    12360    12459: gap of 100 bp
*    12460    14959: contig of 2500 bp in length
*    14960    15059: gap of 100 bp
*    15060    18997: contig of 3938 bp in length
*    18998    19097: gap of 100 bp
*    19098    25489: contig of 6392 bp in length
*    25490    25589: gap of 100 bp
*    25590    31287: contig of 5698 bp in length
*    31288    31387: gap of 100 bp
*    31388    36301: contig of 4914 bp in length
*    36302    36401: gap of 100 bp
*    36402    42560: contig of 6159 bp in length
*    42561    42660: gap of 100 bp
*    42661    48389: contig of 5729 bp in length
*    48390    48489: gap of 100 bp
*    48490    54498: contig of 6009 bp in length
*    54499    54598: gap of 100 bp
*    54599    61031: contig of 6433 bp in length
*    61032    61131: gap of 100 bp
*    61132    68382: contig of 7251 bp in length
*    68383    68482: gap of 100 bp
*    68483    78352: contig of 9870 bp in length
*    78353    78452: gap of 100 bp
*    78453    89588: contig of 11136 bp in length
*    89589    89688: gap of 100 bp
*    89689   100432: contig of 10744 bp in length
*   100433   100532: gap of 100 bp
*   100533   112962: contig of 12430 bp in length
*   112963   113062: gap of 100 bp
*   113063   134930: contig of 21868 bp in length
*   134931   135030: gap of 100 bp
*   135031   165471: contig of 30441 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1. .165471
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /map="22"
                     /clone="RP11-67H19"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1323
                     /note="assembly_fragment"
     misc_feature    1424. .2727
                     /note="assembly_fragment"
     misc_feature    2828. .4344
                     /note="assembly_fragment"
     misc_feature    4445. .5722
                     /note="assembly_fragment"
     misc_feature    5823. .6884
                     /note="assembly_fragment"
     misc_feature    6985. .8290
                     /note="assembly_fragment"
     misc_feature    8391. .9932
                     /note="assembly_fragment"
     misc_feature    10033. .12359
                     /note="assembly_fragment"
     misc_feature    12460. .14959
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     misc_feature    15060. .18997
                     /note="assembly_fragment"
     misc_feature    19098. .25489
                     /note="assembly_fragment"
     misc_feature    25590. .31287
                     /note="assembly_fragment"
     misc_feature    31388. .36301
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    36402. .42560
                     /note="assembly_fragment"
     misc_feature    42661. .48389
                     /note="assembly_fragment"
     misc_feature    48490. .54498
                     /note="assembly_fragment"
     misc_feature    54599. .61031
                     /note="assembly_fragment"
     misc_feature    61132. .68382
                     /note="assembly_fragment"
     misc_feature    68483. .78352
                     /note="assembly_fragment"
     misc_feature    78453. .89588
                     /note="assembly_fragment"
     misc_feature    89689. .100432
                     /note="assembly_fragment"
     misc_feature    100533. .112962
                     /note="assembly_fragment"
     misc_feature    113063. .134930
                     /note="assembly_fragment"
     misc_feature    135031. .165471
                     /note="assembly_fragment"
ORIGIN

  Query Match             21.3%;  Score 2132;  DB 2;  Length 165471;
  Best Local Similarity   96.6%;  Pred. No. 0;
  Matches 2248;  Conservative    0;  Mismatches   55;  Indels   23;  Gaps    6;

Qy          1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGG------AGACTTGCTCTGTTGCCCA 54
              ||||||||||||||| |||||| ||   ||| ||||      || || ||||||||||||
Db      67225 CCATCATATTTCTTTTTTTTTGGGCAAGAGGGGGGAGTCAAAGTCTCGCTCTGTTGCCCA 67166

Qy         55 GGCTGGA--CCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTG 112
              |||||||   |||||| ||||||||||||||||||||||||||||||||||||||||||
```

```
Db      67165 GGCTGGAATGCAGTGGCGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTG 67106
Qy        113 ATTCCCAAATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTT 172
Db      67105 ATTCCCAAGTAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTT 67046
Qy        173 TAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTG 232
Db      67045 TAGCAGATAAGGGGTTTCACCATGTTGGTCAGGCTGGTCTCCAACTCCTGACCTCAAGTG 66986
Qy        233 ATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTC 292
Db      66985 ATCCACCCACTTTGGCTTCCCAAAGCATTGGGAGTACAGGTGTGAGCCACTATGCCCGGC 66926
Qy        293 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGT 352
Db      66925 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTGGAAAGGATGT 66866
Qy        353 CAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAAT 412
Db      66865 CAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACATATTAGTCAAGTTTCTAAAAT 66806
Qy        413 ACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTT 472
Db      66805 ACAGTAATAATGTTGAGAGCTGACGTAGGGACTAACTTAG-----TTTTTGTTTTTTTTT 66751
Qy        473 TTTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAA 532
Db      66750 TTTCAAATTCTCACCTAACTTTGATTTTGCTAAATAAGGACATT-AAAAAAAACCCAAAA 66692
Qy        533 AACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCAT 592
Db      66691 AACTCCACTATTGCATATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCAT 66632
Qy        593 CTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGT 652
Db      66631 CTGCAAGTAGGAAGGGCCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGT 66572
Qy        653 TGTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCC 712
Db      66571 TGTCAGCATTTGTTAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCC 66512
Qy        713 TGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATG 772
Db      66511 TGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATG 66452
Qy        773 TATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAAT 832
Db      66451 TATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAGT 66392
Qy        833 GTGTAGCTCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTA 892
Db      66391 GTGTAGATCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTAGAGTTTA 66332
Qy        893 ACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGT 952
Db      66331 ACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGT 66272
Qy        953 TGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGC 1012
Db      66271 TGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGC 66212
Qy       1013 CCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCT 1072
Db      66211 CCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAGAGACTTCAGGACATTCTTTCCT 66152
Qy       1073 TTACCCTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTT 1132
Db      66151 TTACCCTAATCCTTGATCTTCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTT 66092
Qy       1133 ATTCAAAACTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAG 1192
Db      66091 ATTCAAAACTAAAGATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAG 66032

Qy       1193 ATGTTTCCTTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGA 1252
Db      66031 ATGTTTCCTTGGCTTCTCTGAGCCTACAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGA 65972
Qy       1253 GGGAGGAAAGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCAT 1312
Db      65971 GGGAGGAAAGTAAAGTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCAT 65912
Qy       1313 GCCTTTTGGTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTT 1372
Db      65911 CCCTTTTGGTTTTCTTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTT 65852
Qy       1373 AAAAACAATGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAA 1432
Db      65851 AGAAACAATGCCTATCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAA 65792
Qy       1433 TTCTTCTCCAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAA 1492
Db      65791 TTCTTCTCCAGGAGTGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAA 65732
Qy       1493 TTTTCCACCTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTT 1552
Db      65731 TTTTCCACCTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAAATCAGAGGACCCATTT 65672
Qy       1553 TCTGATGAAAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTAT 1612
Db      65671 TCTGATGAAAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTAT 65612
Qy       1613 ACATTCACAAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCC 1672
Db      65611 ACATTCACAAAGGAGCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCC 65552
Qy       1673 CACCCAGTGCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGA 1732
Db      65551 CACCCAGTGCTTGTTGAAGGCAGAAAGTTGTGGAAAGCAAGCCGGGATTTGAACAGTGGA 65492
Qy       1733 TTGAGGTTTCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGT 1792
Db      65491 TTGAGGTTTCGAATATCCAGTGAACCAAAATATATCAGGGTTCCCGTGGCCAAAATGAGT 65432
Qy       1793 GACCATTCTGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTT 1852
Db      65431 GACCATTCTGAGGTGTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTT 65372
Qy       1853 CTGTGCTCATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTT 1912
Db      65371 CTGTGCTCATTTTGTTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTT 65312
Qy       1913 AGATTTCTTGGATGCAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACA 1972
Db      65311 AGA----TTGGATGCAGATTTGTTTGTATATGAAACAGATTTTAAATTGTTTTGTATACA 65256
Qy       1973 CTGGATTTAAAA-----TAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACA 2027
Db      65255 CTGGATTTAAAATATTTTAGTTTACTAAAGTGTTTTAATGTTTTCATCTTAATCTTCACA 65196
Qy       2028 GTTCTTATAGTCTTTAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTG 2087
Db      65195 GTTCTTATAGTCTTTAGATTTAGGGAGGCTGCTGATGGCATCCACATGTGCATTTTAGTG 65136
Qy       2088 GCATTTAAAATGTATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGA 2147
Db      65135 GCATTTAAAATGTATTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGA 65076
Qy       2148 TTTAAGTCGGTAAAGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAAT 2207
Db      65075 TTTAAGTCGGTAAAGCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAAT 65016
Qy       2208 AAGTGTACTGACTGGATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGC 2267
Db      65015 AAGTGTACTGACTGGATAAAATGCCAGTGATTTAATTAACAAGCACGTTTAACAGGATGC 64956
```

```
Qy       2268 CCTATATATTAGTTAAAAGTGAAGCAATTGAATTAGGTACCTTCTC 2313
              ||||||||||||||||||||||||||||||||||||||| | ||
Db      64955 CCTATATATTAGTTAAAAGTGAAGCAATTGAATTAGGTACCATTTC 64910


RESULT 6
AP000365
LOCUS       AP000365              111123 bp    DNA     linear   PRI 10-JUN-2000
DEFINITION  Homo sapiens genomic DNA, chromosome 22q11.2, clone:KB7G2.
ACCESSION   AP000365 BA000006
VERSION     AP000365.1  GI:5931533
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Riazi,M.A., Brinkman-Mills,P., Johnson,A., Naylor,S.L.,
            Minoshima,S., Shimizu,N., Baldini,A. and McDermid,H.E.
  TITLE     Identification of a putative regulatory subunit of a
            calcium-activated potassium channel in the dup(3q) syndrome region
            and a related sequence on 22q11.2
  JOURNAL   Genomics (1999) In press
REFERENCE   2  (bases 1 to 111123)
  AUTHORS   Shimizu,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUL-1999) Nobuyoshi Shimizu, Keio university, school
            of medicine, Molecular Biology; 35 Shinanomachi, Shinjuku-ku, Tokyo
            160-0016, Japan (E-mail:shimizu@dmb.med.keio.ac.jp,
            Tel:81-3-3351-2370, Fax:81-3-3351-2370)
FEATURES             Location/Qualifiers
     source          1. .111123
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /map="22q11.2"
                     /clone="KB7G2"
     repeat_region   complement(1543. .1601)
                     /rpt_family="MIR"
     repeat_region   2403. .2689
                     /rpt_family="ALU"
     gene            complement(3043. .4982)
                     /gene="KCNMB3L"
     mRNA            complement(join(3043. .3564,4784. .>4982))
                     /gene="KCNMB3L"
                     /pseudo
     exon            complement(3043. .3564)
                     /gene="KCNMB3L"
                     /number=3
                     /pseudo
     exon            complement(4784. .4982)
                     /gene="KCNMB3L"
                     /number=2
                     /pseudo
     repeat_region   complement(5071. .5224)
                     /rpt_family="ALU"
     repeat_region   5399. .5475
                     /rpt_family="ALU"
     repeat_region   5752. .6041
                     /rpt_family="ALU"
     repeat_region   6074. .6355
                     /rpt_family="ALU"
     repeat_region   6690. .6989
                     /rpt_family="ALU"
     repeat_region   7282. .7541
                     /rpt_family="ALU"
     repeat_region   complement(7574. .7861)
                     /rpt_family="ALU"
     repeat_region   8728. .8768
                     /rpt_family="ALU"
     repeat_region   complement(8974. .9150)
                     /rpt_family="ALU"
     repeat_region   complement(9200. .9457)
                     /rpt_family="ALU"
     repeat_region   9735. .10034
                     /rpt_family="ALU"
     repeat_region   10336. .10918
                     /rpt_family="ALU"
     repeat_region   11000. .11287
                     /rpt_family="ALU"
     repeat_region   complement(11694. .12297)
                     /rpt_family="ALU"
     repeat_region   12699. .12989
                     /rpt_family="ALU"
     gene            complement(13110. .15070)
                     /gene="CESK1"
     mRNA            complement(13110. .15070)
                     /gene="CESK1"
                     /note="single exon gene"
     CDS             complement(13154. .14827)
                     /gene="CESK1"
                     /note="single exon gene"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAA84680.1"
                     /db_xref="GI:5931534"
                     /translation="MDSTVPSALELPQRLALNPRESPRSPEEEEPHLLSSLAAVQTLA
                     SVIRPCYGPHGRQKFLVTMKGETVCTGCATAILRALELEHPAAWLLREAGQTQAENSG
                     DGTAFVVLLTEALLEQAEQLLKAGLPRPQLREAYATATAEVLATLPSLAIQSLGPLED
                     PSWALHSVMNTHTLSPMDHLTKLVAHACWAIKELDGSFKPERVGVCALPGGTLEDSCL
                     LPGLAISGKLCGQMATVLSGARVALFACPFGPAHPNAPATARLSSPADLAQFSKGSDQ
                     LLEKQVGQLAAAGINVAVVLGEVDEETLTLADKYGIVVIQARSWMEIIYLSEVLDTPL
                     LPRLLPPQRPGKCQRVYRQELGDGLAVVFEWECTGTPALTVVLRGATTQGLRSAEQAV
                     YHGIDAYFQLCQDPRLIPGAGATEMALAKMLSDKGSRLEGPSGPAFLAFAWALKYLPK
                     TLAENAGLAVSDVMAEMSGVHQGGNLLMGVGTEGIINVAQEGVWDTLIVKAQGFRAVA
                     EVVLQLVTVDEIVVAKKSPTHQEIWNPDSKKTKKHPPPVETKKILGLNN"
     repeat_region   complement(15324. .15882)
                     /rpt_family="ALU"
     repeat_region   16430. .16579
                     /rpt_family="ALU"
     repeat_region   complement(16772. .17108)
                     /rpt_family="ALU"
     repeat_region   18110. .18279
                     /rpt_family="ALU"
     repeat_region   18353. .18613
                     /rpt_family="ALU"
     repeat_region   complement(19159. .19425)
                     /rpt_family="ALU"
     repeat_region   complement(19669. .19969)
                     /rpt_family="ALU"
     repeat_region   complement(20241. .20832)
                     /rpt_family="ALU"
     repeat_region   20544. .20598
                     /rpt_family="ALU"
     repeat_region   complement(20669. .21107)
                     /rpt_family="ALU"
     repeat_region   complement(26952. .27229)
                     /rpt_family="ALU"
     repeat_region   complement(28472. .28770)
                     /rpt_family="ALU"
     repeat_region   29345. .29608
                     /rpt_family="ALU"
     repeat_region   complement(29969. .30260)
                     /rpt_family="ALU"
     repeat_region   30347. .30641
                     /rpt_family="ALU"
     repeat_region   33585. .34024
                     /rpt_family="LTR7"
     repeat_region   36986. .37419
                     /rpt_family="LTR7"
     repeat_region   complement(37429. .37677)
                     /rpt_family="ALU"
     repeat_region   complement(40282. .40473)
                     /rpt_family="L1"
     repeat_region   complement(41057. .41745)
```

```
                     /rpt_family="L1"
     repeat_region   46673. .46956
                     /rpt_family="ALU"
     repeat_region   complement(47515. .48025)
                     /rpt_family="MER42"
     repeat_region   49234. .49276
                     /rpt_family="THE1"
     repeat_region   complement(52666. .52717)
                     /rpt_family="MLT1"
     repeat_region   53759. .54035
                     /rpt_family="ALU"
     repeat_region   complement(55174. .55236)
                     /rpt_family="ALU"
     repeat_region   complement(55782. .56814)
                     /rpt_family="L1"
     repeat_region   59423. .59715
                     /rpt_family="ALU"
     repeat_region   61070. .61335
                     /rpt_family="ALU"
     repeat_region   complement(61766. .62009)
                     /rpt_family="MER43"
     repeat_region   complement(63650. .67174)
                     /rpt_family="L1"
     repeat_region   67188. .67468
                     /rpt_family="ALU"
     repeat_region   complement(70783. .71062)
                     /rpt_family="ALU"
     repeat_region   72529. .72783
                     /rpt_family="MER7"
     repeat_region   complement(74296. .74439)
                     /rpt_family="MER5"
     repeat_region   75366. .75612
                     /rpt_family="ALU"
     repeat_region   75666. .75943
                     /rpt_family="ALU"
     repeat_region   78172. .78402
                     /rpt_family="ALU"
     repeat_region   complement(78597. .79032)
                     /rpt_family="ALU"
     repeat_region   79128. .79408
                     /rpt_family="ALU"
     repeat_region   79926. .80200
                     /rpt_family="ALU"
     repeat_region   82675. .89455
                     /rpt_family="L1"
     repeat_region   92817. .93048
                     /rpt_family="ALU"
     repeat_region   complement(93239. .93306)
                     /rpt_family="MER46"
     repeat_region   complement(94031. .94324)
                     /rpt_family="ALU"
     repeat_region   complement(94710. .94999)
                     /rpt_family="ALU"
     repeat_region   complement(99945. .100244)
                     /rpt_family="ALU"
     repeat_region   100754. .101034
                     /rpt_family="ALU"
     repeat_region   complement(104640. .104845)
                     /rpt_family="ALU"
     gene            complement(106967. .111123)
                     /gene="F8VWPP"

  Query Match             20.2%;  Score 2015;  DB 9;  Length 111123;
  Best Local Similarity   95.3%;  Pred. No. 0;
  Matches 2178;  Conservative    0;  Mismatches   60;  Indels   47;  Gaps    8;

Qy         36 AGACTTGCTCTGTTGCCCAGGCTGGA--CCAGTGGTGCGATCTTGGCTCACTGCAACCTC 93
              || || || |||||||||||||||||    ||||||  |||||||||||||||| |||||
Db      37429 AGTCTCGCCCTGTTGCCCAGGCTGGAATGCAGTGGCACGATCTTGGCTCACTGCTACCTC 37488

Qy         94 CACCTCCTGGGTTCAAGTGATTCCCAAATAGCTGGGATTACAGGTGTGTATTACCATGCC 153
              ||| |||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      37489 CACTTCCTGGGTTCAAGTGATTCCCAAGTAGCTGGGATTACAGGTGTGTATTACCATGCC 37548

Qy        154 CAGCTAATTTTTGTATTTTTAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTC 213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37549 CAGCTAATTTTTGTATTTTTAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTC 37608

Qy        214 CAACTCCTGGCCTCATGTGATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTATAGGT 273
              |||||||||||||| |||||||||||||||||||| |||||||||||||||||| |||
Db      37609 CAACTCCTGGCCTCAAGTGATCCACCCACTTCGGCTACCCAAAGCATTGGGAGTACAGGG 37668

Qy        274 GTGAGCCACTATACCCGTCCTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGT 333
              ||||||||||||| |||  ||||||||||||||||||||||||||||||||||||||||||||
Db      37669 GTGAGCCACTATGCCCAGCCTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGT 37728

Qy        334 CTCTTTTTCTAAAGGATGTCAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACGTA 393
              |||||||| |||||||||||||||||||| |||||||||||||||||||||||||||| ||
Db      37729 CTCTTTTTGGAAAGGATGTCAGTAGAGAAATGGAGTTCCCCAAAATTACAGTTTCACATA 37788

Qy        394 TTAGTCAAGTTTCTAAAATACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTGGT 453
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      37789 TTAGTAAAGTTTCTAAAATACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTAG- 37847

Qy        454 TTTTTTTTTTTTTTTTTTTTTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGAC 513
              |||||||||||||||||||||| ||||  ||||||||||||||||||||||||||||||
Db      37848 ------TTTTTTTTTTTTTTTTCAAATTGTCACCTAACTTTGATTTTGCTAAATAAGGAC 37901

Qy        514 ATTAAAAAAAAAAACCAAAAAACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAA 573
              ||| |||| |||||||||||| |||||||||| ||||||||||||||||||||||||||||
Db      37902 ATT-AAAAGAAAACCAAAAAAGTCCACTATTGGCTATTGCCACTATTTGATTTTTTAAAA 37960

Qy        574 AATAAGCGTATTTTAGCATCTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCT 633
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37961 AATAAGCGTATTTTAGCATCTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCT 38020

Qy        634 TGGCCAGGGAAAACAGCGTTGTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTT 693
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      38021 TGGCCAGGGAAAACAGCGTTGTCAGCATTTGATAACTGTTTTTCTAGGGTATGTGCTGTT 38080

Qy        694 ATTCAGTTAAAACCTTGCCTGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAA 753
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db      38081 ATTCAGTTAAAACCTTGCCTGGGACGCTAGCATTCGGTAAATACTTGTTGAATAAGCAAA 38140

Qy        754 TGAAACTTAAGCTTCTATGTATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTA 813
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38141 TGAAACTTAAGCTTCTATGTATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTA 38200

Qy        814 ATTGAATATTCTTTTCAATGTGTAGCTCTATCCCCAGAACCACAGAATATTGGAACTGTA 873
              |||||||||||||||| |||||||| ||||| ||||||||||||||||||||||||||||
Db      38201 ATTGAATATTCTTTTCAGTGTGTAGATCTATACCCAGAACCACAGAATATTGGAACTGTA 38260

Qy        874 AAGGCCATCCTATAGTTTAACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTG 933
              ||                ||||||||| ||||||               ||||||||||||||
Db      38261 AA---------------AACCAACTGCATTAAAT-----------AGATGTGGTATGTG 38293

Qy        934 GCAGTGACAACTTGAAGGTTGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCA 993
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      38294 GCAGTGACAACTTGAAGGTTGTGACTAGAACTCCGGTCTCTGGAGTGTTCTATTATATCA 38353

Qy        994 CACCAAGCTGGTCACCAGCCCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGA 1053
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38354 CACCAAGCTGGTCACCAGCCCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGA 38413

Qy       1054 CTTCAGGACATTCTTTCCTTTACCCTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTT 1113
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      38414 CTTCAGGACATTCTTTCCTTTACCCTAATCCTTGATCTTCAGTCTTATTTAGAAAAGCTT 38473

Qy       1114 AATGTTAAAGATCTAGTTTATTCAAAACTAAAGATAACAAGGAGTATGAGAATTTCTATT 1173
              |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db      38474 AATGTTAAAGATCTAGTTTATTCAAAACTAAAGAT---AAGGAGTATGAGAATTTCTATT 38530

Qy       1174 TCGGAGTGTAAAGGAGGAGATGTTTCCTTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCT 1233
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      38531 TCGGAGTGTAAAGGAGGAGATGTTTCCTTGGCTTCTCTGAGCCTACAGGCCTTCCTTGCT 38590
```

```
Qy       1234 CTTTAAGGAAGTAGAGAGAGGGAGGAAAGTAAAGTATGCTTTTGTTTTTTAAGGTTACTT 1293
             |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db      38591 CTTTAAGGAAGTAGAGAGAGAGGAAAGTAAAGTATGCTTTTGTTTTTTAAGGTTACTT 38650

Qy       1294 TGCTGGGAGTAGTTTGCATGCCTTTTGGTTTTCTTGGGTGGAATTAACTGACTTAAGTTT 1353
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      38651 TGCTGGGAGTAGTTTGCATGCCATTTGGTTTTCTTGGGTGGAATTAACTGACTTAAGTTT 38710

Qy       1354 TAAGTAGTTGGGACTATTTAAAAACAATGCCTATCCAATGTTTGCCATAAAGGCAGAGGG 1413
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      38711 TAAGTAGTTGGGACTATTTAGAAACAATGCCTATCCAATGTTTGCCATAAAGGCAGAGGG 38770

Qy       1414 TATTGGCTTTAGAAGTTAATTCTTCTCCAGGAGTGAAAATTAGCTTCTAAACCAGAAGCA 1473
             |||||||||||||||||||||||||||||||||||| |||| ||||||||||||||||||
Db      38771 TATTGGCTTTAGAAGTTAATTCTTCTCCAGGAGTGAAAATGAGCTGCTAAACCAGAAGCA 38830

Qy       1474 GCAGAGCTAAATAAAGTAATTTTCCACCTGGCCAGTGCATGATGTGAAAGGTAGATTAAA 1533
             |||||||||| |||||||||||||||||||||||||||||||||||||||||||| ||||
Db      38831 GCAGAGCTAAAGAAAGTAATTTTCCACCTGGCCAGTGCATGATGTGAAAGGTAGACTAAA 38890

Qy       1534 AAAATGAGAGGGCCCATTTTCTGATGAAAGACTAAGCCATGTTGAAACAGCCCTGTTGAG 1593
             | ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db      38891 AGAATGAGAGGACCCATTTTCTGATGAAAGACTAAGCCATGTTGAAACAGCCCTGTTGAG 38950

Qy       1594 GATTTTATTTTAAATCTATACATTCACAAAGGAGCTTTGTGTATGTCTTTCCCTATTTGT 1653
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38951 GATTTTATTTTAAATCTATACATTCACAAAGGAGCTTTGTGTATGTCTTTCCCTATTTGT 39010

Qy       1654 TGTTTGGACTAGGAAGCCCCACCCAGTGCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAG 1713
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      39011 TGTTTGGACTAGGAAGCCCCACCCAGTGCTTGTTGAAGGCAGAAAGTTGTTGAAAGCAAT 39070

Qy       1714 CTGGGATTTGAACAGTGGATTGAGGTTTCGAATATCCAGTGAACCAAAATATATCAGGGT 1773
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39071 CCGGGATTTGAACAGTGGATTGAGGTTTCGAATATCCAGTGAACCAAAATATATCAGGGT 39130

Qy       1774 TCCCCTGGCCAAGATGAGTGACCATTCTGAGGTGTTAAGTATTTCTTGAATGGGGATTTT 1833
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db      39131 TCCCCTGGCCAAGATGAGTGACCATTCTGAGGTGTTACGTATTTCTTGAATGGGGATTTT 39190

Qy       1834 AGGAAAAGTTTCTGTATTTCTGTGCTCATTTTGTTGACCTCTGTATGTGCAAAATCTCTA 1893
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39191 AGGAAAAGTTTCTGTATTTCTGTGCTCATTTTGTTGACCTCTGTATGTGCAAAATCTCTA 39250

Qy       1894 AGGGGGTGTTTGGGCACTTAGATTTCTTGGATGCAGATTTGTTTGTATATGAAACAAATT 1953
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39251 AGGGGGTGTTTGGGCACTTAGATTTCTTGGATGCAGATTTGTTTGTGTATGAAACAAATT 39310

Qy       1954 TTAAATTGTTTTGTATACACTGGATTTAAAA-----TAGTTTACTAAAGTGTTTTAATTT 2008
             |||||||||| |||||||||||||||||||     |||||||||||||||||||||||| |
Db      39311 TTAAATTGTTTCGTATACACTGGATTTAAAATAGTTTAGTTTACTAAAGTGTTTTAATGT 39370

Qy       2009 TTTCATCTTAATTTTCACAGTTCTTATAGTCTTTAGATTTAGGGAGGCTGTTGATGGCAT 2068
             |||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db      39371 TTTCATCTTAATTTT--CAGTTCTTATAGTCTTTAGATTTAGGGAGGCTGTTGATGGCAT 39428

Qy       2069 CCACATGTGCATTTTAGTGGCATTTAAAATGTATTCAGCTGAATTTAACAATTTCTGACC 2128
             |||||| ||||||||| || |||||||||||||||| ||||||||||||||||||||||||
Db      39429 CCACATATGCATTTTAATGTCATTTAAAATGTATTCTGCTGAATTTAACAATTTCTGACC 39488

Qy       2129 TAAAACTTGACATTTTAGATTTAAGTCGGTAAAGCACTGATTTAAACTGGATTTTAACTG 2188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39489 TAAAACTTGACATTTTAGATTTAAGTCGGTAAAGCACTGATTTAAACTGGATTTTAACTG 39548

Qy       2189 GATGAAATTCTGATTTAATAAGTGTACTGACTGGATAAAATGCCAATGATTTAATTAACA 2248
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      39549 GATGAAATTCTGATTTAATGAGTGTACTGACTGGATAAAATGCCAATGATTTAATTAACA 39608

Qy       2249 AGCACGTTTAACAGGATGCCCTATATATTAGTTAAAAGTGAAGCAATTGAATTAGGTACC 2308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39609 AGCACGTTTAACAGGATGCCCTATATATTAGTTAAAAGTGAAGCAATTGAATTAGGTACC 39668

Qy       2309 TTCTC 2313
              | ||
Db      39669 ATTTC 39673
```

RESULT 7
AC141274/c

```
LOCUS       AC141274              187590 bp    DNA     linear   HTG 11-MAR-2003
DEFINITION  Homo sapiens chromosome 16 clone RP11-164D11, WORKING DRAFT
            SEQUENCE, 6 unordered pieces.
ACCESSION   AC141274
VERSION     AC141274.1  GI:28913054
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 187590)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 187590)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2003) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 469424
            Center clone name: RPCI-11_164D11
            -------
            Summary Statistics
            Consensus quality: 184633 bases at least Q40
            Consensus quality: 185350 bases at least Q30
            Consensus quality: 185738 bases at least Q20
            Estimated insert size: 175000; agarose-fp estimation
            Estimated insert size: 187090; sum-of-contigs estimation
            Quality coverage:  12.44 in Q20 bases; agarose-fp estimation
            Quality coverage:  11.63 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1783: contig of 1783 bp in length
            *     1784     1883: gap of unknown length
            *     1884     4310: contig of 2427 bp in length
            *     4311     4410: gap of unknown length
            *     4411    39416: contig of 35006 bp in length
            *    39417    39516: gap of unknown length
            *    39517    80057: contig of 40541 bp in length
            *    80058    80157: gap of unknown length
            *    80158   116056: contig of 35899 bp in length
            *   116057   116156: gap of unknown length
            *   116157   187590: contig of 71434 bp in length.
FEATURES             Location/Qualifiers
     source          1. .187590
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="RP11-164D11"
                     /clone_lib="RPCI human BAC library 11"
ORIGIN
```

```
  Query Match             20.2%;  Score 2015;  DB 2;  Length 187590;
  Best Local Similarity   95.3%;  Pred. No. 0;
```

```
Matches 2178;  Conservative    0;  Mismatches   60;  Indels   47;  Gaps    8;

Qy         36 AGACTTGCTCTGTTGCCCAGGCTGGA--CCAGTGGTGCGATCTTGGCTCACTGCAACCTC  93
Db     174350 AGTCTCGCCCTGTTGCCCAGGCTGGAATGCAGTGGCACGATCTTGGCTCACTGCTACCTC 174291

Qy         94 CACCTCCTGGGTTCAAGTGATTCCCAAATAGCTGGGATTACAGGTGTGTATTACCATGCC 153
Db     174290 CACTTCCTGGGTTCAAGTGATTCCCAAGTAGCTGGGATTACAGGTGTGTATTACCATGCC 174231

Qy        154 CAGCTAATTTTTGTATTTTTAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTC 213
Db     174230 CAGCTAATTTTTGTATTTTTAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTC 174171

Qy        214 CAACTCCTGGCCTCATGTGATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTATAGGT 273
Db     174170 CAACTCCTGGCCTCAAGTGATCCACCCACTTCGGCTACCCAAAGCATTGGGAGTACAGGG 174111

Qy        274 GTGAGCCACTATACCCGTCCTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGT 333
Db     174110 GTGAGCCACTATGCCCAGCCTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGT 174051

Qy        334 CTCTTTTTCTAAAGGATGTCAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACGTA 393
Db     174050 CTCTTTTTGGAAAGGATGTCAGTAGAGAAATGGAGTTCCCCAAAATTACAGTTTCACATA 173991

Qy        394 TTAGTCAAGTTTCTAAAATACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTGGT 453
Db     173990 TTAGTAAAGTTTCTAAAATACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTAG- 173932

Qy        454 TTTTTTTTTTTTTTTTTTTTTTTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGAC 513
Db     173931 ------TTTTTTTTTTTTTTTTCAAATTGTCACCTAACTTTGATTTTGCTAAATAAGGAC 173878

Qy        514 ATTAAAAAAAAAAACCAAAAAACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAA 573
Db     173877 ATT-AAAAGAAAACCAAAAAAGTCCACTATTGGCTATTGCCACTATTTGATTTTTTAAAA 173819

Qy        574 AATAAGCGTATTTTAGCATCTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCT 633
Db     173818 AATAAGCGTATTTTAGCATCTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCT 173759

Qy        634 TGGCCAGGGAAAACAGCGTTGTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTT 693
Db     173758 TGGCCAGGGAAAACAGCGTTGTCAGCATTTGATAACTGTTTTTCTAGGGTATGTGCTGTT 173699

Qy        694 ATTCAGTTAAAACCTTGCCTGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAA 753
Db     173698 ATTCAGTTAAAACCTTGCCTGGGACGCTAGCATTCGGTAAATACTTGTTGAATAAGCAAA 173639

Qy        754 TGAAACTTAAGCTTCTATGTATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTA 813
Db     173638 TGAAACTTAAGCTTCTATGTATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTA 173579

Qy        814 ATTGAATATTCTTTTCAATGTGTAGCTCTATCCCCAGAACCACAGAATATTGGAACTGTA 873
Db     173578 ATTGAATATTCTTTTCAGTGTGTAGATCTATACCCAGAACCACAGAATATTGGAACTGTA 173519

Qy        874 AAGGCCATCCTATAGTTTAACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTG 933
Db     173518 AA----------------AACCAACTGCATTAAAT-----------AGATGTGGTATGTG 173486

Qy        934 GCAGTGACAACTTGAAGGTTGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCA 993
Db     173485 GCAGTGACAACTTGAAGGTTGTGACTAGAACTCCGGTCTCTGGAGTGTTCTATTATATCA 173426

Qy        994 CACCAAGCTGGTCACCAGCCCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGA 1053
Db     173425 CACCAAGCTGGTCACCAGCCCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGA 173366

Qy       1054 CTTCAGGACATTCTTTCCTTTACCCTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTT 1113
Db     173365 CTTCAGGACATTCTTTCCTTTACCCTAATCCTTGATCTTCAGTCTTATTTAGAAAAGCTT 173306

Qy       1114 AATGTTAAAGATCTAGTTTATTCAAAACTAAAGATAACAAGGAGTATGAGAATTTCTATT 1173
Db     173305 AATGTTAAAGATCTAGTTTATTCAAAACTAAAGAT---AAGGAGTATGAGAATTTCTATT 173249

Qy       1174 TCGGAGTGTAAAGGAGGAGATGTTTCCTTGGCTTCTCTGAGCCTGCAGGCCTTCCTTGCT 1233
Db     173248 TCGGAGTGTAAAGGAGGAGATGTTTCCTTGGCTTCTCTGAGCCTACAGGCCTTCCTTGCT 173189

Qy       1234 CTTTAAGGAAGTAGAGAGAGGGAGGAAAGTAAAGTATGCTTTTGTTTTTTAAGGTTACTT 1293
Db     173188 CTTTAAGGAAGTAGAGAGAGAGGAAAGTAAAGTATGCTTTTGTTTTTTAAGGTTACTT 173129

Qy       1294 TGCTGGGAGTAGTTTGCATGCCTTTTGGTTTTCTTGGGTGGAATTAACTGACTTAAGTTT 1353
Db     173128 TGCTGGGAGTAGTTTGCATGCCATTTGGTTTTCTTGGGTGGAATTAACTGACTTAAGTTT 173069

Qy       1354 TAAGTAGTTGGGACTATTTAAAAACAATGCCTATCCAATGTTTGCCATAAAGGCAGAGGG 1413
Db     173068 TAAGTAGTTGGGACTATTTAGAAACAATGCCTATCCAATGTTTGCCATAAAGGCAGAGGG 173009

Qy       1414 TATTGGCTTTAGAAGTTAATTCTTCTCCAGGAGTGAAAATTAGCTTCTAAACCAGAAGCA 1473
Db     173008 TATTGGCTTTAGAAGTTAATTCTTCTCCAGGAGTGAAAATGAGCTGCTAAACCAGAAGCA 172949

Qy       1474 GCAGAGCTAAATAAAGTAATTTTCCACCTGGCCAGTGCATGATGTGAAAGGTAGATTAAA 1533
Db     172948 GCAGAGCTAAAGAAAGTAATTTTCCACCTGGCCAGTGCATGATGTGAAAGGTAGACTAAA 172889

Qy       1534 AAAATGAGAGGGCCCATTTTCTGATGAAAGACTAAGCCATGTTGAAACAGCCCTGTTGAG 1593
Db     172888 AGAATGAGAGGACCCATTTTCTGATGAAAGACTAAGCCATGTTGAAACAGCCCTGTTGAG 172829

Qy       1594 GATTTTATTTTAAATCTATACATTCACAAAGGAGCTTTGTGTATGTCTTTCCCTATTTGT 1653
Db     172828 GATTTTATTTTAAATCTATACATTCACAAAGGAGCTTTGTGTATGTCTTTCCCTATTTGT 172769

Qy       1654 TGTTTGGACTAGGAAGCCCCACCCAGTGCTTGTTGAAGGCAGAAAGTCGTTGAAAGCAAG 1713
Db     172768 TGTTTGGACTAGGAAGCCCCACCCAGTGCTTGTTGAAGGCAGAAAGTTGTTGAAAGCAAT 172709

Qy       1714 CTGGGATTTGAACAGTGGATTGAGGTTTCGAATATCCAGTGAACCAAAAATATATCAGGGT 1773
Db     172708 CCGGGATTTGAACAGTGGATTGAGGTTTCGAATATCCAGTGAACCAAAAATATATCAGGGT 172649

Qy       1774 TCCCCTGGCCAAGATGAGTGACCATTCTGAGGTGTTAAGTATTTCTTGAATGGGGATTTT 1833
Db     172648 TCCCCTGGCCAAGATGAGTGACCATTCTGAGGTGTTACGTATTTCTTGAATGGGGATTTT 172589

Qy       1834 AGGAAAAGTTTCTGTATTTCTGTGCTCATTTTGTTGACCTCTGTATGTGCAAAATCTCTA 1893
Db     172588 AGGAAAAGTTTCTGTATTTCTGTGCTCATTTTGTTGACCTCTGTATGTGCAAAATCTCTA 172529

Qy       1894 AGGGGGGTGTTTGGGCACTTAGATTTCTTGGATGCAGATTTGTTTGTATATGAAACAAATT 1953
Db     172528 AGGGGGTGTTTGGGCACTTAGATTTCTTGGATGCAGATTTGTTTGTGTATGAAACAAATT 172469

Qy       1954 TTAAATTGTTTTGTATACACTGGATTTAAAA-----TAGTTTACTAAAGTGTTTTAATTT 2008
Db     172468 TTAAATTGTTTCGTATACACTGGATTTAAAATAGTTTAGTTTACTAAAGTGTTTTAATGT 172409

Qy       2009 TTTCATCTTAATTTTCACAGTTCTTATAGTCTTTAGATTTAGGGAGGCTGTTGATGGCAT 2068
Db     172408 TTTCATCTTAATTTT--CAGTTCTTATAGTCTTTAGATTTAGGGAGGCTGTTGATGGCAT 172351

Qy       2069 CCACATGTGCATTTTAGTGGCATTTAAAATGTATTCAGCTGAATTTAACAATTTCTGACC 2128
Db     172350 CCACATATGCATTTTAATGTCATTTAAAATGTATTCTGCTGAATTTAACAATTTCTGACC 172291

Qy       2129 TAAAACTTGACATTTTAGATTTAAGTCGGTAAAGCACTGATTTAAACTGGATTTTAACTG 2188
Db     172290 TAAAACTTGACATTTTAGATTTAAGTCGGTAAAGCACTGATTTAAACTGGATTTTAACTG 172231
```

```
Qy       2189 GATGAAATTCTGATTTAATAAGTGTACTGACTGGATAAAATGCCAATGATTTAATTAACA 2248
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     172230 GATGAAATTCTGATTTAATGAGTGTACTGACTGGATAAAATGCCAATGATTTAATTAACA 172171

Qy       2249 AGCACGTTTAACAGGATGCCCTATATATTAGTTAAAAGTGAAGCAATTGAATTAGGTACC 2308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     172170 AGCACGTTTAACAGGATGCCCTATATATTAGTTAAAAGTGAAGCAATTGAATTAGGTACC 172111

Qy       2309 TTCTC 2313
                | ||
Db     172110 ATTTC 172106


RESULT 8
AR157086
LOCUS       AR157086               5632 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 3 from patent US 6242590.
ACCESSION   AR157086
VERSION     AR157086.1  GI:15125790
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 5632)
  AUTHORS   Cowsert,L.M.
  TITLE     Antisense modulation of zinc finger protein-217 expression
  JOURNAL   Patent: US 6242590-A 3 05-JUN-2001;
FEATURES             Location/Qualifiers
     source          1. .5632
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             16.4%;  Score 1639.4;  DB 6;  Length 5632;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1640;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 2533
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 60

Qy       2534 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 2593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 120

Qy       2594 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 2653
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 180

Qy       2654 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 2713
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 240

Qy       2714 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300

Qy       2774 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 2833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 360

Qy       2834 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 2893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 420

Qy       2894 TACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 2953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 480

Qy       2954 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 540

Qy       3014 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600

Qy       3074 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 660

Qy       3134 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 3193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 720

Qy       3194 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 3253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 780

Qy       3254 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 3313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 840

Qy       3314 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 900

Qy       3374 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 960

Qy       3434 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 1020

Qy       3494 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 1080

Qy       3554 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 3613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 1140

Qy       3614 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 3673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 1200

Qy       3674 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 3733
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 1260

Qy       3734 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 3793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 1320

Qy       3794 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 3853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 1380

Qy       3854 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 3913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 1440

Qy       3914 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 3973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 1500

Qy       3974 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCCTCTGGATGA 4033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCCTCTGGATGA 1560

Qy       4034 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 4093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 1620

Qy       4094 CGAAGGAATCCATCTGGGTAA 4114
```

```
                    |||||||||||||||||| |||
Db       1621 CGAAGGAATCCATCTGGATAA 1641


RESULT 9
AF041259
LOCUS       AF041259               5632 bp    mRNA    linear   PRI 22-JUL-1998
DEFINITION  Homo sapiens breast cancer putative transcription factor (ZABC1)
            mRNA, complete cds.
ACCESSION   AF041259
VERSION     AF041259.1  GI:3335396
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5632)
  AUTHORS   Collins,C., Rommens,J.M., Kowbel,D., Godfrey,T., Tanner,M.,
            Hwang,S.-I., Polikoff,D., Nonet,G., Cochran,J., Myambo,K.,
            Jay,K.E., Froula,J., Cloutier,T., Kuo,W.-L., Yaswen,P., Dairkee,S.,
            Giovanola,J., Hutchinson,G.B., Isola,J., Kallioniemi,O.-P.,
            Palazzolo,M., Martin,C., Ericsson,C., Pinkel,D., Albertson,D.,
            Li,W.-B. and Gray,J.W.
  TITLE     Positional cloning of ZNF217 and NABC1: genes amplified at 20q13.2
            and overexpressed in breast carcinoma
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 95 (15), 8703-8708 (1998)
  MEDLINE   98337979
   PUBMED   9671742
REFERENCE   2  (bases 1 to 5632)
  AUTHORS   Collins,C., Rommens,J.M., Kowbel,D., Godfrey,T., Tanner,M.,
            Hwang,S.-I., Polikoff,D., Nonet,G., Cochran,J., Myambo,K.,
            Jay,K.E., Froula,J., Cloutier,T., Kuo,W.-L., Yaswen,P., Dairkee,S.,
            Giovanola,J., Hutchinson,G.B., Isola,J., Kallioniemi,O.-P.,
            Palazzolo,M., Martin,C., Ericsson,C., Pinkel,D., Albertson,D.,
            Li,W.-B. and Gray,J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-1998) Life Sciences, Lawrence Berkeley National
            Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
FEATURES             Location/Qualifiers
     source          1. .5632
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20q13.2-q13.3"
     gene            1. .5632
                     /gene="ZABC1"
     CDS             272. .3418
                     /gene="ZABC1"
                     /note="Kruppel-like"
                     /codon_start=1
                     /product="breast cancer putative transcription factor"
                     /protein_id="AAC39895.1"
                     /db_xref="GI:3335397"
                     /translation="MQSKVTGNMPTQSLLMYMDGPEVIGSSLGSPMEMEDALSMKGTA
                     VVPFRATQEKNVIQIEGYMPLDCMFCSQTFTHSEDLNKHVLMQHRPTLCEPAVLRVEA
                     EYLSPLDKSQVRTEPPKEKNCKENEFSCEVCGQTFRVAFDVEIHMRTHKDSFTYGCNM
                     CGRRFKEPWFLKNHMRTHNGKSGARSKLQQGLESSPATINEVVQVHAAESISSPYKIC
                     MVCGFLFPNKESLIEHRKVHTKKTAFGTSSAQTDSPQGGMPSSREDFLQLFNLRPKSH
                     PETGKKPVRCIPQLDPFTTFQAWQLATKGKVAICQEVKESGQEGSTDNDDSSSEKELG
                     ETNKGSCAGLSQEKEKCKHSHGEAPSVDADPKLPSSKEKPTHCSECGKAFRTYHQLVL
                     HSRVHKKDRRAGAESPTMSVDGRQPGTCSPDLAAPLDENGAVDRGEGGSEDGSEDGLP
                     EGIHLDKNDDGGKIKHLTSSRECSYCGKFFRSNYYLNIHLRTHTGEKPYKCEFCEYAA
                     AQKTSLRYHLERHHKEKQTDVAAEVKNDGKNQDTEDALLTADSAQTKNLKRFFDGAKD
                     VTGSPPAKQLKEMPSVFQNVLGSAVLSPAHKDTQDFHKNAADDSADKVNKNPTPAYLD
                     LLKKRSAVETQANNLICRTKADVTPPPDGSTTHNLEVSPKEKQTETAADCRYRPSVDC
                     HEKPLNLSVGALHNCPAISLSKSLIPSITCPFCTFKTFYPEVLMMHQRLEHKYNPDVH
                     KNCRNKSLLRSRRTGCPPALLGKDVPPLSSFCKPKPKSAFPAQSKSLPSAKGKQSPPG
                     PGKAPLTSGIDSSTLAPSNLKSHRPQQNVGVQGAATRQQQSEMFPKTSVSPAPDKTKR
                     PETKLKPLPVAPSQPTLGSSNINGSIDYPAKNDSPWAPPGRDYFCNRSASNTAAEFGE
                     PLPKRLKSSVVALDVDQPGANYRRGYDLPKYHMVRGITSLLPQDCVYPSQALPPKPRF
                     LSSSEVDSPNVLTVQKPYGGSGPLYTCVPAGSPASSSTLEGKRPVSYQHLSNSMAQKR
                     NYENFIGNAHYRPNDKKT"
     misc_feature    470. .535
                     /gene="ZABC1"
                     /note="zinc finger 1"
     misc_feature    659. .721
                     /gene="ZABC1"
                     /note="zinc finger 2"
     misc_feature    744. .805
                     /gene="ZABC1"
                     /note="zinc finger 3"
     misc_feature    923. .985
                     /gene="ZABC1"
                     /note="zinc finger 4"
     misc_feature    1400. .1463
                     /gene="ZABC1"
                     /note="zinc finger 5"
     misc_feature    1688. .1750
                     /gene="ZABC1"
                     /note="zinc finger 6"
     misc_feature    1772. .1834
                     /gene="ZABC1"
                     /note="zinc finger 7"
     misc_feature    2405. .2470
                     /gene="ZABC1"
                     /note="zinc finger 8"
ORIGIN

  Query Match             16.4%;  Score 1639.4;  DB 9;  Length 5632;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1640;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 2533
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 60

Qy       2534 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 2593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 120

Qy       2594 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 2653
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 180

Qy       2654 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 2713
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 240

Qy       2714 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300

Qy       2774 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 2833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 360

Qy       2834 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 2893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 420

Qy       2894 TACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 2953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 480

Qy       2954 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 540

Qy       3014 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600

Qy       3074 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 660
```

```
Qy       3134 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 3193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 720

Qy       3194 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 3253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 780

Qy       3254 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 3313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 840

Qy       3314 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 900

Qy       3374 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 960

Qy       3434 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 1020

Qy       3494 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 1080

Qy       3554 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 3613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 1140

Qy       3614 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 3673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 1200

Qy       3674 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 3733
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 1260

Qy       3734 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 3793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 1320

Qy       3794 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 3853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 1380

Qy       3854 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 3913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 1440

Qy       3914 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 3973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 1500

Qy       3974 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 4033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 1560

Qy       4034 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 4093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 1620

Qy       4094 CGAAGGAATCCATCTGGGTAA 4114
             |||||||||||||||| |||
Db        1621 CGAAGGAATCCATCTGGATAA 1641


RESULT 10
CQ726295
LOCUS       CQ726295               5633 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 12229 from Patent WO02068579.
ACCESSION   CQ726295
VERSION     CQ726295.1  GI:42288937
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 12229 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .5633
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             16.4%;  Score 1639.4;  DB 6;  Length 5633;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1640;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 2533
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTG 60

Qy       2534 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 2593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATGCCTTGTTTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAG 120

Qy       2594 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 2653
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTT 180

Qy       2654 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 2713
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTC 240

Qy       2714 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CAGAAATTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300

Qy       2774 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 2833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AACTCAATCCCTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAG 360

Qy       2834 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 2893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TCCGATGGAGATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGC 420

Qy       2894 TACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 2953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 TACACAAGAAAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTG 480

Qy       2954 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CAGCCAGACCTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCC 540

Qy       3014 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TACCCTCTGTGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600

Qy       3074 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AAGTCAAGTGCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTG 660

Qy       3134 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 3193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 720
```

```
Qy       3194 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 3253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTT 780

Qy       3254 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 3313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTTAAAAATCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCA 840

Qy       3314 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAG 900

Qy       3374 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CATCTCCTCTCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAG 960

Qy       3434 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 1020

Qy       3494 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 GACAGACTCTCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAA 1080

Qy       3554 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 3613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 CTTGAGACCAAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCT 1140

Qy       3614 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 3673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CGATCCGTTCACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTG 1200

Qy       3674 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 3733
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGA 1260

Qy       3734 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 3793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAA 1320

Qy       3794 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 3853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 GTGCAAACACTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAG 1380

Qy       3854 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 3913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 CAAGGAGAAGCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCT 1440

Qy       3914 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 3973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 GGTCTTGCACTCCAGGGTCCACAAGAAGGACCGGAGGGCCGGCGCGGAGTCGCCCACCAT 1500

Qy       3974 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 4033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 GTCTGTGGACGGGAGGCAGCCGGGGACGTGTTCTCCTGACCTCGCCGCCCCTCTGGATGA 1560

Qy       4034 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 4093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 AAATGGAGCCGTGGATCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 1620

Qy       4094 CGAAGGAATCCATCTGGGTAA 4114
             |||||||||||||||| |||
Db        1621 CGAAGGAATCCATCTGGATAA 1641
```

RESULT 11
BD085734
LOCUS BD085734 3186 bp DNA linear PAT 27-AUG-2002
DEFINITION Genes from the 20q13 amplicon and their uses.
ACCESSION BD085734
VERSION BD085734.1 GI:22631344
KEYWORDS JP 2001524802-A/10.

SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 3186)
AUTHORS Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and Rommens,J.
TITLE Genes from the 20q13 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 10 04-DEC-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
COMMENT OS Artificial Sequence
PN JP 2001524802-A/10
PD 04-DEC-2001
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR 17-JAN-1997 US 08/785532
PI JOE W GRAY,COLIN CONRAD COLLINS,SOO IN HWANG,TONY GODFREY, PI DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11,C12Q1/68,A61K48/00
CC Description of Artificial Sequence:ZABC1 Open Reading Frame FH
Key Location/Qualifiers
FT source 1. .3186
FT /organism='Artificial Sequence'.
FEATURES Location/Qualifiers
source 1. .3186
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN

Query Match 15.6%; Score 1559.6; DB 6; Length 3186;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy       8286 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACATCTCT 8345
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1482 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAGAAGACATCTCT 1541

Qy       8346 GAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 8405
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1542 GAGGTATCACTTGGAGAGACATCACAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 1601

Qy       8406 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCGCAAAC 8465
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1602 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAACCGCTGACAGTGCGCAAAC 1661

Qy       8466 CAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCTGCAAA 8525
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1662 CAAAAATTTGAAAAGATTTTTTGATGGTGCCAAAGATGTTACAGGCAGTCCACCTGCAAA 1721

Qy       8526 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTCTCACC 8585
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1722 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAATGTTCTGGGCAGCGCTGTCCTCTCACC 1781

Qy       8586 AGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGATAAAGT 8645
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1782 AGCACACAAAGATACTCAGGATTTCCATAAAAATGCAGCTGATGACAGTGCTGATAAAGT 1841

Qy       8646 GAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAC 8705
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1842 GAATAAAAACCCTACCCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAC 1901

Qy       8706 TCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGATGGCAG 8765
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1902 TCAGGCAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCCTCCTCCGGATGGCAG 1961

Qy       8766 TACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCTGACTG 8825
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1962 TACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAAGCAAACGGAGACCGCAGCTGACTG 2021

Qy       8826 CAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCT 8885
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2022 CAGATACAGGCCAAGTGTGGATTGTCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCT 2081
```

```
Qy       8886 TCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGTCCATT 8945
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2082 TCACAATTGCCCGGCAATTTCTTTGAGTAAAAGTTTGATTCCAAGTATCACCTGTCCATT 2141

Qy       8946 TTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAGCATAA 9005
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2142 TTGTACCTTCAAGACATTTTATCCAGAAGTTTTAATGATGCACCAGAGACTGGAGCATAA 2201

Qy       9006 ATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGACGTAC 9065
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2202 ATACAATCCTGACGTTCATAAAAACTGTCGAAACAAGTCCTTGCTTAGAAGTCGACGTAC 2261

Qy       9066 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCCCTAGTTTCTGTAAACC 9125
             ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       2262 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCTCCCCTCTCTAGTTTCTGTAAACC 2321

Qy       9126 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 9185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2322 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 2381

Qy       9186 GAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 9245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2382 GAGCCCTCCTGGGCCAGGCAAGGCCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 2441

Qy       9246 CCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCCGCCAC 9305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2442 CCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAATGTGGGGGTCCAAGGGGCCGCCAC 2501

Qy       9306 CAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGATAAGAC 9365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2502 CAGGCAACAGCAATCTGAGATGTTTCCTAAAACCAGTGTTTCCCCTGCACCGGATAAGAC 2561

Qy       9366 AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACCCTCGG 9425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2562 AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCCACCCTCGG 2621

Qy       9426 CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGGGCACC 9485
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2622 CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAGAACGACAGCCCGTGGGCACC 2681

Qy       9486 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTTGGTGA 9545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2682 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGCAGCAGAATTTGGTGA 2741

Qy       9546 GCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCCGGGGC 9605
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2742 GCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAGCCCGGGGC 2801

Qy       9606 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT 9665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2802 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT 2861

Qy       9666 GTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTCCTGAG 9725
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2862 GTTACCGCAGGACTGTGTGTATCCGTCGCAGGCGCTGCCTCCCAAACCAAGGTTCCTGAG 2921

Qy       9726 CTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGCTCCGG 9785
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2922 CTCCAGCGAGGTCGATTCTCCAAATGTGCTGACTGTTCAGAAGCCCTATGGTGGCTCCGG 2981

Qy       9786 GCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTAT 9845
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       2982 GCCACTTTACACTTGTGTGCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTCT 3041

Qy       9846 TGCATGAGGGGCGT 9859
             ||   || |   ||
Db       3042 TGGTGGATGTCAGT 3055
```

```
RESULT 12
AR070326
LOCUS       AR070326               1507 bp    DNA     linear   PAT 18-FEB-2000
DEFINITION  Sequence 3 from patent US 5892010.
ACCESSION   AR070326
VERSION     AR070326.1  GI:7221214
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1507)
  AUTHORS   Gray,J., Collins,C., Hwang,S.-i., Godfrey,T., Kowbel,D. and
            Rommens,J.
  TITLE     Genes from the 20Q13 amplicon and their uses
  JOURNAL   Patent: US 5892010-A 3 06-APR-1999;
FEATURES             Location/Qualifiers
     source          1. .1507
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             13.9%;  Score 1387.2;  DB 6;  Length 1507;
  Best Local Similarity   97.3%;  Pred. No. 4.7e-311;
  Matches 1470;  Conservative    3;  Mismatches   32;  Indels    6;  Gaps    6;

Qy       2424 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGA 2483
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGC 61

Qy       2484 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGT 2543
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGT 121

Qy       2544 TTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGG 2603
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGG 181

Qy       2604 ACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTTGGAACCCAGG 2663
             ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        182 ACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGG-TTTTGGAACCCAGG 240

Qy       2664 GCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTCCAGAAATTGC 2723
             |||||||| |||||||||||| ||||||| |||||||||| ||||||||||||||||||||
Db        241 ACTGTTAGA-GGGTGAGTGACAGGTCTTAC-AGTGGCCTTAATCCAACTCCAGAAATTGC 298

Qy       2724 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 2783
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 358

Qy       2784 CTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAG 2843
             |||||||||||||||||| ||||| ||||||||||||||||||| ||||||||||||||||
Db        359 CTCTTAATGTACATGGAT-GGCCAAGAGTGATTGGCAGCTCTCTTGCCAGTCCGATGGAG 417

Qy       2844 ATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGAA 2903
             ||||| |||||||||||||||||| || ||||||| | |          |||||
Db        418 ATGGA-GATGCCTTGTCAATGAAAGGGCCCNCTGTTGTCAATTCCGAGCTACACAAAGAA 476

Qy       2904 AAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 2962
             ||||||||| |||| |||||||||| ||||||||||||||||||||||||||||||||||||
Db        477 AAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 536

Qy       2963 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 3022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 596

Qy       3023 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 3082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 656

Qy       3083 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 3142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 716

Qy       3143 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 3202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 776
```

```
Qy       3203 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 3262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 836

Qy       3263 TCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 3322
             ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCACATGCGGACRCATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 896

Qy       3323 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTC 3382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTC 956

Qy       3383 TCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 3442
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 TCCTTGCAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 1016

Qy       3443 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTC 3502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTC 1076

Qy       3503 TCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGACC 3562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 TCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGACC 1136

Qy       3563 AAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGTT 3622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1137 AAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGTT 1196

Qy       3623 CACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTGCCAAGAAGT 3682
             ||||||||||||||||||:||||||||||||||:||||||||||||||||||||||||||||
Db       1197 CACCACCTTCCAGGCTTGGCAKCTGGCTACCAAAGGAAWAGTTGCCATTTGCCAAGAAGT 1256

Qy       3683 GAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGAGAAGGAGCT 3742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1257 GAAGGAATTGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGAGAAGGAGCT 1316

Qy       3743 TGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACA 3802
             |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db       1317 TGGAGAAACAAATAAGAACCATTGTGCAGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACA 1376

Qy       3803 CTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAA 3862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1377 CTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAA 1436

Qy       3863 GCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCTGGTCTTGCA 3922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1437 GCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCTGGTCTTGCA 1496

Qy       3923 CTCCAGGGTCC 3933
             ||||||||||||
Db       1497 CTCCAGGGTCC 1507
```

```
RESULT 13
BD085727
LOCUS       BD085727               1507 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Genes from the 20q13 amplicon and their uses.
ACCESSION   BD085727
VERSION     BD085727.1  GI:22631337
KEYWORDS    JP 2001524802-A/3.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 1507)
  AUTHORS   Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and
            Rommens,J.
  TITLE     Genes from the 20q13 amplicon and their uses
  JOURNAL   Patent: JP 2001524802-A 3 04-DEC-2001;
            THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
COMMENT     OS   Artificial Sequence
            PN   JP 2001524802-A/3
            PD   04-DEC-2001
            PF   15-JUL-1997 JP 1998506264
            PR   15-JUL-1996 US    08/680395,16-OCT-1996 US    08/731499 PR
            17-JAN-1997 US    08/785532
            PI   JOE W GRAY,COLIN CONRAD COLLINS,SOO IN HWANG,TONY GODFREY, PI
             DAVID KOWBEL,
            PI   JOHANNA ROMMENS
            PC   C12N15/11,C12Q1/68,A61K48/00
            CC   Description of Artificial Sequence:cc49
            CC   N is A, G, G, T, or U
            FH   Key             Location/Qualifiers
            FT   modified base   (447).
FEATURES            Location/Qualifiers
     source         1. .1507
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
ORIGIN
```

```
  Query Match           13.9%;  Score 1387.2;  DB 6;  Length 1507;
  Best Local Similarity  97.3%;  Pred. No. 4.7e-311;
  Matches 1470;  Conservative    3;  Mismatches   32;  Indels    6;  Gaps    6;

Qy       2424 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGA 2483
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGC 61

Qy       2484 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGT 2543
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCTGTATGCCTTGT 121

Qy       2544 TTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGG 2603
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TTGTGGAAAGAACCAGTGACACCATCACTGAGCTTCCTAAAAGTTCGAAGAAGTTAGAGG 181

Qy       2604 ACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGGTTTTTGGAACCCAGG 2663
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        182 ACTATACACTTTCTTTTGAACTTTTATAATAAATATTTGCTCTGG-TTTTGGAACCCAGG 240

Qy       2664 GCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCTTATTCCAACTCCAGAAATTGC 2723
             |||||||| |||||||||||| ||||||| |||||||||| ||||||||||||||||||||
Db        241 ACTGTTAGA-GGGTGAGTGACAGGTCTTAC-AGTGGCCTTAATCCAACTCCAGAAATTGC 298

Qy       2724 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 2783
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 358

Qy       2784 CTCTTAATGTACATGGATGGGCCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAG 2843
             |||||||||||||||||| ||||| |||||||||||||||||| ||||||||||||||||
Db        359 CTCTTAATGTACATGGAT-GGCCAAGAGTGATTGGCAGCTCTCTTGCCAGTCCGATGGAG 417

Qy       2844 ATGGAGGATGCCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCCGAGCTACACAAGAA 2903
             ||||| |||||||||||||||||||| || |||||||    | |  ' |||||||||||||
Db        418 ATGGA-GATGCCTTGTCAATGAAAGGGCCCNCTGTTGTCAATTCCGAGCTACACAAAGAA 476

Qy       2904 AAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 2962
             ||||||||| |||| |||||||||| ||||||||||||||||||||||||||||||||||||
Db        477 AAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 536

Qy       2963 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 3022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 CTTCACACATTCAGAAGACCTTAATAAACATGTCTTAATGCAACACCGGCCTACCCTCTG 596

Qy       3023 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 3082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 TGAACCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 656

Qy       3083 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 3142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 GCGAACAGAACCTCCCAAGGAAAAGAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATG 716

Qy       3143 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 3202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTC 776
```

```
Qy       3203 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 3262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 836

Qy       3263 TCACATGCGGACACATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 3322
             ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCACATGCGGACRCATAATGGCAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTTGGA 896

Qy       3323 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTC 3382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 GAGTAGTCCAGCAACGATCAACGAGGTCGTCCAGGTGCACGCGGCCGAGAGCATCTCCTC 956

Qy       3383 TCCTTACAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 3442
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 TCCTTGCAAAATCTGCATGGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 1016

Qy       3443 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTC 3502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTC 1076

Qy       3503 TCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGACC 3562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 TCCACAAGGAGGAATGCCGTCCTCGAGGGAGGACTTCCTGCAGTTGTTCAACTTGAGACC 1136

Qy       3563 AAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGTT 3622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1137 AAAATCTCACCCTGAAACGGGGAAGAAGCCTGTCAGATGCATCCCTCAGCTCGATCCGTT 1196

Qy       3623 CACCACCTTCCAGGCTTGGCAGCTGGCTACCAAAGGAAAAGTTGCCATTTGCCAAGAAGT 3682
             ||||||||||||||||||||:||||||||||||||:||||||||||||||||||||||||
Db       1197 CACCACCTTCCAGGCTTGGCAKCTGGCTACCAAAGGAAWAGTTGCCATTTGCCAAGAAGT 1256

Qy       3683 GAAGGAATCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGAGAAGGAGCT 3742
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1257 GAAGGAATTGGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCCGAGAAGGAGCT 1316

Qy       3743 TGGAGAAACAAATAAGGGCAGTTGTGCAGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACA 3802
             ||||||||||||||||  |  ||||||||||||||||||||||||||||||||||||||||
Db       1317 TGGAGAAACAAATAAGAACCATTGTGCAGGCCTCTCGCAAGAGAAAGAGAAGTGCAAACA 1376

Qy       3803 CTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAA 3862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1377 CTCCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAA 1436

Qy       3863 GCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCTGGTCTTGCA 3922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1437 GCCCACTCACTGCTCCGAGTGCGGCAAAGCTTTCAGAACCTACCACCAGCTGGTCTTGCA 1496

Qy       3923 CTCCAGGGTCC 3933
             ||||||||||||
Db       1497 CTCCAGGGTCC 1507
```

```
RESULT 14
HS449O17/c
LOCUS       HS449O17              59586 bp    DNA     linear   PRI 04-MAR-2003
DEFINITION  Human DNA sequence from clone RP3-449O17 on chromosome
            22q13.1-13.2, complete sequence.
ACCESSION   Z98749
VERSION     Z98749.11  GI:4775611
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 59586)
  AUTHORS   Lloyd,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: cloneрequest@sanger.ac.uk
COMMENT     On May 11, 1999 this sequence version replaced gi:4688875.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep --------------
            Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            This sequence was generated from part of bacterial clone contigs of
            human chromosome 22, constructed by the Sanger Centre Chromosome 22
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr22
            RP3-449O17 is from the library RPCI-3 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2.
FEATURES             Location/Qualifiers
     source          1. .59586
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /map="q13.1-13.2"
                     /clone="RP3-449O17"
                     /clone_lib="RPCI-3"
ORIGIN
```

```
  Query Match             12.2%;  Score 1223.6;  DB 9;  Length 59586;
  Best Local Similarity   97.8%;  Pred. No. 8.4e-273;
  Matches 1264;  Conservative    0;  Mismatches   19;  Indels    9;  Gaps    2;

Qy       1027 TCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCTTTACCCTAATCCTT 1086
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      59586 TCCATTGTGATAGCAACAAAGAAGAGACTTCAGGACATTCTTTCCTTTACCCTAATCCTT 59527

Qy       1087 GATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAACTAAAG 1146
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      59526 GATCTTCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTCAAAACTAAAG 59467

Qy       1147 ATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCCTTGGCT 1206
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      59466 ATAACAAGGAGTATGAGAATTTCTATTTCGGAGTGTAAAGGAGGAGATGTTTCCTTGGCT 59407

Qy       1207 TCTCTGAGCCTGCAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAAAGTAAA 1266
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db      59406 TCTCTGAGCCTACAGGCCTTCCTTGCTCTTTAAGGAAGTAGAGAGAGGGAGGAAAGTAAA 59347

Qy       1267 GTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATGCCTTTTGGTTTTC 1326
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      59346 GTATGCTTTTGTTTTTTAAGGTTACTTTGCTGGGAGTAGTTTGCATCCCTTTTGGTTTTC 59287

Qy       1327 TTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAAAAACAATGCCTA 1386
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      59286 TTGGGTGGAATTAACTGACTTAAGTTTTAAGTAGTTGGGACTATTTAGAAACAATGCCTA 59227

Qy       1387 TCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTCCAGGAG 1446
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        59226 TCCAATGTTTGCCATAAAGGCAGAGGGTATTGGCTTTAGAAGTTAATTCTTCTCCAGGAG 59167

Qy         1447 TGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCACCTGGCC 1506
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        59166 TGAAAATTAGCTTCTAAACCAGAAGCAGCAGAGCTAAATAAAGTAATTTTCCACCTGGCC 59107

Qy         1507 AGTGCATGATGTGAAAGGTAGATTAAAAAAATGAGAGGGCCCATTTTCTGATGAAAGACT 1566
                |||||||||||||||||||||||||||||||| ||||| |||||||||||||||||||||
Db        59106 AGTGCATGATGTGAAAGGTAGATTAAAAAAATCAGAGGACCCATTTTCTGATGAAAGACT 59047

Qy         1567 AAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTAAATCTATACATTCACAAAGGA 1626
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        59046 AAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTAAATCTATACATTCACAAAGGA 58987

Qy         1627 GCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGTGCTTGT 1686
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        58986 GCTTTGTGTATGTCTTTCCCTATTTGTTGTTTGGACTAGGAAGCCCCACCCAGTGCTTGT 58927

Qy         1687 TGAAGGCAGAAAGTCGTTGAAAGCAAGCTGGGATTTGAACAGTGGATTGAGGTTTCGAAT 1746
                ||||||||||||| || ||||||||| |||||||||||||||||||||||||||||||||
Db        58926 TGAAGGCAGAAAGTTGTGGAAAGCAAGCCGGGATTTGAACAGTGGATTGAGGTTTCGAAT 58867

Qy         1747 ATCCAGTGAACCAAAATATATCAGGGTTCCCCTGGCCAAGATGAGTGACCATTCTGAGGT 1806
                |||||||||||||||||||||||||||||||| ||||||| |||||||||||||||||||
Db        58866 ATCCAGTGAACCAAAATATATCAGGGTTCCCGTGGCCAAAATGAGTGACCATTCTGAGGT 58807

Qy         1807 GTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTCATTTTG 1866
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        58806 GTTAAGTATTTCTTGAATGGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTCATTTTG 58747

Qy         1867 TTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGATTTCTTGGATG 1926
                ||||||||||||||||||||||||||||||||||||||||||||||||    ||||||||
Db        58746 TTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTTTGGGCACTTAGA----TTGGATG 58691

Qy         1927 CAGATTTGTTTGTATATGAAACAAATTTTAAATTGTTTTGTATACACTGGATTTAAAA-- 1984
                |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        58690 CAGATTTGTTTGTATATGAAACAGATTTTAAATTGTTTTGTATACACTGGATTTAAAATA 58631

Qy         1985 ---TAGTTTACTAAAGTGTTTTAATTTTTTCATCTTAATTTTCACAGTTCTTATAGTCTT 2041
                   ||||||||||||||||||||| |||||||||||| ||||||||||||||||||||||
Db        58630 TTTTAGTTTACTAAAGTGTTTTAATGTTTTCATCTTAATCTTCACAGTTCTTATAGTCTT 58571

Qy         2042 TAGATTTAGGGAGGCTGTTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGTA 2101
                ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        58570 TAGATTTAGGGAGGCTGCTGATGGCATCCACATGTGCATTTTAGTGGCATTTAAAATGTA 58511

Qy         2102 TTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAAA 2161
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        58510 TTCAGCTGAATTTAACAATTTCTGACCTAAAACTTGACATTTTAGATTTAAGTCGGTAAA 58451

Qy         2162 GCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACTG 2221
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        58450 GCACTGATTTAAACTGGATTTTAACTGGATGAAATTCTGATTTAATAAGTGTACTGACTG 58391

Qy         2222 GATAAAATGCCAATGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGTT 2281
                |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db        58390 GATAAAATGCCAGTGATTTAATTAACAAGCACGTTTAACAGGATGCCCTATATATTAGTT 58331

Qy         2282 AAAAGTGAAGCAATTGAATTAGGTACCTTCTC 2313
                ||||||||||||||||||||||||||| | ||
Db        58330 AAAAGTGAAGCAATTGAATTAGGTACCATTTC 58299
```

```
RESULT 15
HS434P1/c
LOCUS       HS434P1               124990 bp    DNA     linear   PRI 04-MAR-2003
DEFINITION  Human DNA sequence from clone RP3-434P1 on chromosome 22, complete
            sequence.
ACCESSION   Z97056
VERSION     Z97056.1  GI:2832593
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 124990)
  AUTHORS   Connor,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 5, 1998 this sequence version replaced gi:2462403.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep -------------
            Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            This sequence was generated from part of bacterial clone contigs of
            human chromosome 22, constructed by the Sanger Centre Chromosome 22
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr22
            RP3-434P1 is from the library RPCI-3 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2.
FEATURES             Location/Qualifiers
     source          1..124990
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704P01434"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /clone="RP3-434P1"
                     /clone_lib="RPCI-3"
ORIGIN

  Query Match             10.1%;  Score 1006.8;  DB 9;  Length 124990;
  Best Local Similarity   95.5%;  Pred. No. 2.3e-222;
  Matches 1083;  Conservative    0;  Mismatches   37;  Indels   14;  Gaps    4;

Qy            1 CCATCATATTTCTTATTTTTTTGGGCGGAGAGGGG------AGACTTGCTCTGTTGCCCA 54
                ||||||||||||| |||||| ||   ||| ||||      || || |||||||||||||
Db         1128 CCATCATATTTCTTTTTTTTTGGGCAAGAGGGGGGAGTCAAAGTCTCGCTCTGTTGCCCA 1069

Qy           55 GGCTGGA--CCAGTGGTGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTG 112
                |||||||  |||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         1068 GGCTGGAATGCAGTGGCGCGATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTG 1009

Qy          113 ATTCCCAAATAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTT 172
                |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1008 ATTCCCAAGTAGCTGGGATTACAGGTGTGTATTACCATGCCCAGCTAATTTTTGTATTTT 949

Qy          173 TAGCAGATAAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTG 232
                ||||||||||||||||||||||||||||| |||||||||||||||||||| ||||| |||
Db          948 TAGCAGATAAGGGGTTTCACCATGTTGGTCAGGCTGGTCTCCAACTCCTGACCTCAAGTG 889
```

```
Qy        233 ATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTATAGGTGTGAGCCACTATACCCGTC 292
              |||||||||||||||||||||||||||||||||||| |||||||||||||||| |||| |
Db        888 ATCCACCCACTTCGGCTTCCCAAAGCATTGGGAGTACAGGTGTGAGCCACTATGCCCGGC 829

Qy        293 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTCTAAAGGATGT 352
              |||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db        828 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTTGGAAAGGATGT 769

Qy        353 CAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACGTATTAGTCAAGTTTCTAAAAT 412
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        768 CAGTAGAGAAGTGGAGTTCCCCAAAATTACAGTTTCACATATTAGTCAAGTTTCTAAAAT 709

Qy        413 ACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTGGTTTTTTTTTTTTTTTTTTTT 472
              |||||||||||||||||||||||| |||||||||||| |     ||||| |||||||||
Db        708 ACAGTAATAATGTTGAGAGCTGACGTAGGGACTAACTTAG-----TTTTTGTTTTTTTTT 654

Qy        473 TTTCAAATTCTCACTGAACTTTGATTTTGCTAAATAAGGACATTAAAAAAAAAAACCAAAA 532
              |||||||||||||  |||||||||||||||||||||||||||| |||||||| ||||||
Db        653 TTTCAAATTCTCACCTAACTTTGATTTTGCTAAATAAGGACATT-AAAAAAAACCCAAAA 595

Qy        533 AACTCCACTATTGCCTATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCAT 592
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        594 AACTCCACTATTGCATATTGCCACTATTTGATTTTTTAAAAAATAAGCGTATTTTAGCAT 535

Qy        593 CTAAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGT 652
              ||  |||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        534 CTGCAAGTAGGAAGGGCCTCAAATAAATGAGTCTTTGTTCTTGGCCAGGGAAAACAGCGT 475

Qy        653 TGTCAGAATTTGATAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCC 712
              |||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        474 TGTCAGCATTTGTTAACTGTTTTTCTAGGGTATGTGCTGTTATTCAGTTAAAACCTTGCC 415

Qy        713 TGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATG 772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TGGGACGCTAGCATTCAGTAAATACTTGTTGAATAAGCAAATGAAACTTAAGCTTCTATG 355

Qy        773 TATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAAT 832
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        354 TATAGAAACCTAAGTCACTTCACATTCTGATTAGCAGAGTAATTGAATATTCTTTTCAGT 295

Qy        833 GTGTAGCTCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTATAGTTTA 892
              |||||| |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        294 GTGTAGATCTATCCCCAGAACCACAGAATATTGGAACTGTAAAGGCCATCCTAGAGTTTA 235

Qy        893 ACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGT 952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 ACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGT 175

Qy        953 TGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGC 1012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 TGTGACTAGAACTCGGGTCTCTGGAGTGTTCTATTATATCACACCAAGCTGGTCACCAGC 115

Qy       1013 CCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAAAGACTTCAGGACATTCTTTCCT 1072
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        114 CCATGTGTTGATCCTCCATTGTGATAGCAACAAAGAAGAGACTTCAGGACATTCTTTCCT 55

Qy       1073 TTACCCTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATC 1126
              |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db         54 TTACCCTAATCCTTGATCTTCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATC 1

Search completed: July 31, 2005, 20:22:46
Job time : 27419 secs
```